Deep Learning Lab

Assignment 3

Università della Svizzera Italiana

Caselli Alex

November 2019

Indice

1 Report

1.1 Text preprocessing

The text is splitted into a list of characters and then converted to lower case.

Number of unique characters = 104

Characters	Frequencies
'space'	420427
e	258701
t	180210
a	165255
o	156942
i	142186
n	137458
...	...
@	1

Complete list of frequencies available in the Appendix (subsection 2.1)
And then each character becomes **identifiable by an intege**r which is **its position** in the unique characters list. To reduce noise in the text **i have removed the first 185 lines** of the text which were just a **chapters index** without any useful information for the learning procedure.

1.2 Batches generation

To train our network with that big text it have to **split it into batches**, to accomplish this task the text is split in block of length equals to the batch size and every **block is also split in sub sequences** usually of the same length (it could happen that the last batch has shorter sub sequences) such that **batch i contains the i-th sub sequence of each block**. if the last batch has sub sequences of length shorter that 2, the batch is discarded because is not possible to define inputs and targets
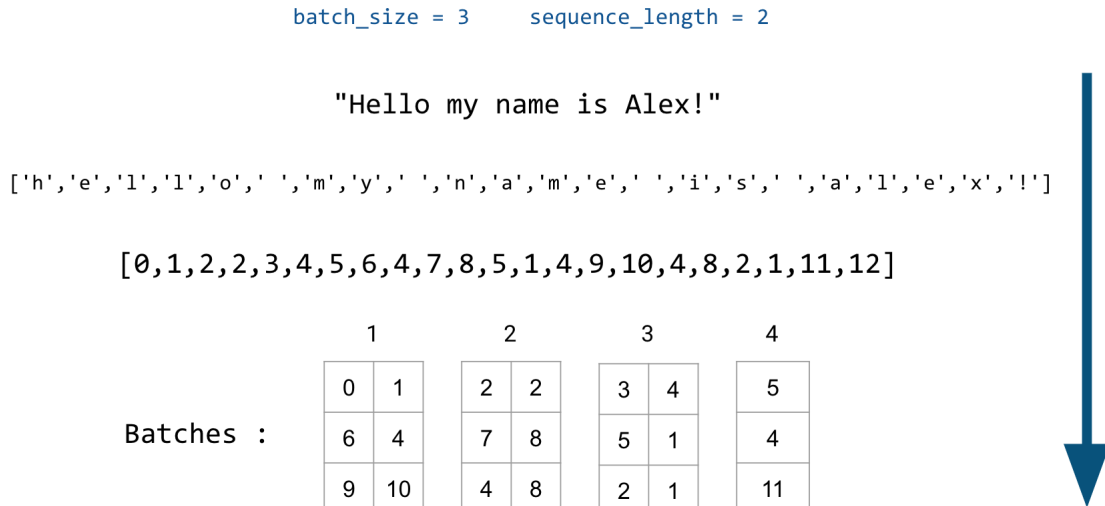
```
batch_size = 3      sequence_length = 2


         "Hello my name is Alex!"

['h','e','l','l','o',' ','m','y',' ','n','a','m','e',' ','i','s',' ','a','l','e','x','!']

     [0,1,2,2,3,4,5,6,4,7,8,5,1,4,9,10,4,8,2,1,11,12]

              1        2        3        4

            ┌──┬──┐ ┌──┬──┐ ┌──┬──┐ ┌────┐
            │0 │1 │ │2 │2 │ │3 │4 │ │5   │
            ├──┼──┤ ├──┼──┤ ├──┼──┤ ├────┤
  Batches : │6 │4 │ │7 │8 │ │5 │1 │ │4   │
            ├──┼──┤ ├──┼──┤ ├──┼──┤ ├────┤
            │9 │10│ │4 │8 │ │2 │1 │ │11  │
            └──┴──┘ └──┴──┘ └──┴──┘ └────┘
```

Figura 1: Example of the method used

1.3   Training

Using this hyperparameters:

- Learning Rate: $1e - 2 = 10^{-2}$

- Batch size: 16

- Subsequences Length: 256

- Number of epochs: 5

- Number of layers: 2

- Number of LSTM units: 256

I have trained the first net using the training loss to trace the performances at each epoch averaging the loss obtained for each batch.

To allow the network preserving the internal state and make it able to learn long sequences of characters the **internal state has been substituted by a Placeholder**, whose value is a **zero state for the first batch** of epochs and the **final state of the previous batch for the next batches** of each epoch.

The **state is a tuple**, so for any state assignation the state tensor has to be converted to a list of tuples The **inputs and the targets for each batch is defined like this** to avoid any problem with shorter sub sequences:

Inputs X = batch[:, :batch.shape[1]-1]

Inputs Y = batch[:, 1:]

And also the sequence_length parameter for the RNN is equal to batch.shape[1]
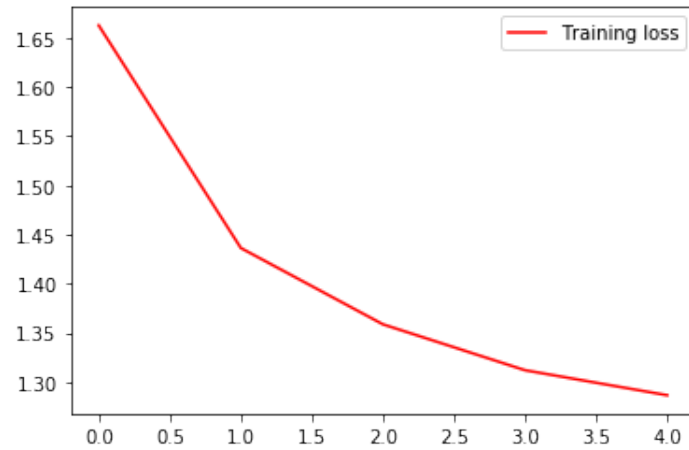
Figura 2: Train loss over the 5 epochs with the suggested hyperparameters

From the plot we can see how the model reduce the loss nicely obtaining a **loss value around 1.25 at the last epoch**, but what if we train it for 10 epochs?
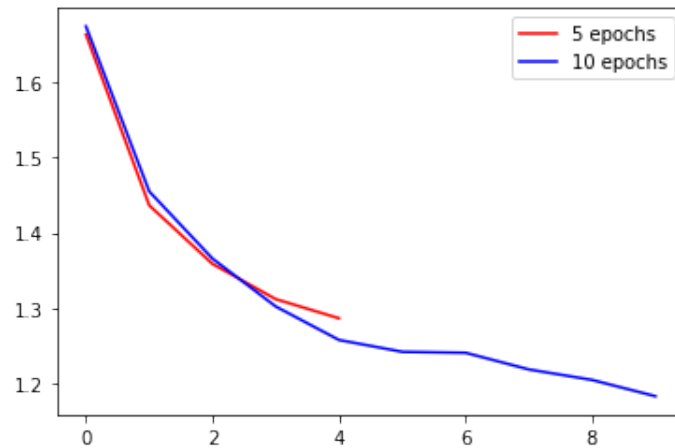

Figura 3: Train loss over 5 and 10 epochs with the suggested hyperparameters

We can see that with more epochs the **loss continues to decrease but without a validation set we cannot say if we are now overfitting or not**, but we can generate some sentences from the 2 models and try to see if one is better than the other (not an objective measure).

1.4   Text Generation

Lets now generate 20 sentences of 256 characters each and **pick the best 5 for each model to compare them**.

Model 1 (5 epochs)
Sentence 1

Ring for a gilt with matered avenger fixed toward himself the usual blossicsaw of each, as i would have been od vuscolding, then, that madse used aroused a probability, whose confidence, paletia, had strawled with me sharp."

while hereobsenfide the count g

Sentence 2

Jus, well, which, he was becomes, by the ground, had striking to him. the count is flashed during the garden."

"mademoiselle pupling at him impress," said the god.

"you?" asked his fight a sight of the paniting france of thought.

"my own; do not poor gla

Sentence 3

E m. de morcerf. there he wished desire."

"perhaps i must be more like poperesmed piazza do myself—ali bied, his gun, as well here, remained italy, i must see him for as really never take post, or my wish you prepared for uneasires, but to him. 1817," repl

Sentence 4

Gevered solicititious quiets, are permitted in secures. at this tumature, who even before him. assist us dia of a low same truth of the dridious some way the apartment, how congucted by the kind.

40140m

"you must killed me you, it is a modest gotit what

Sentence 5

Grarm, situated, which abaud he had contained so esfored his place of minute than sufficied been disappeared or tigited report, his massteld sound of a place in, and was his reason alone of his bad dark. these seconders renders with the carriage let ubtered

Model 2 (10 epochs)
Sentence 1

During the priest, "such complexion was proved obeund the royal. strange gold handsolum, haydée," replied the two bankfredian imaginant, entered that the banker. turning to the fenguance promised, and each, with dophing under the village, ransing to see th

Sentence 2

E my first. you will keep chimn for back one of the black."

"for his heart of you, ha pissol, accept, unfortunate person?"

"i shall love monte cristo?"

"yes, what?"

"you say you could come intending any more or caradiaro, you cannot be a ground, will go

Sentence 3

Ouve where what did disappear specison he alter. the count had just pressing him wed and a vessel, who, bicklyoughliano delick; the trie. albert's are such right, time cickned to agition of their uncernof the second upilio alighted throug" north of a few na

Sentence 4

Ty, attonions of the injulice of unendarence after the strange movements, and monti, and the dark of his questions which he had his most two broverners be killed. "i have entirely an armmortors commitsed; to inform me of his weeding the thut the pobsal, i h

Sentence 5

Rigle hands by see the tect. albert gave her humure with the round of unuses, madame_' avinionions of hunting more skyless a young suntal from their mind of him, while caderousse, "had, if he know, and gave him

As we can see there is almost no difference between the sentences generated by the 2 models (at least to me, because it is not an objective valutation). For this reason we can use just 5 epochs which require less time to train the network, but to be sure i have also trained the network for 20 epochs but even if the results was quite better (less single characters, less random new lines) it required a very long time to get trained (you can find 5 example sentences at the end of the report (**subsection 2.3**).
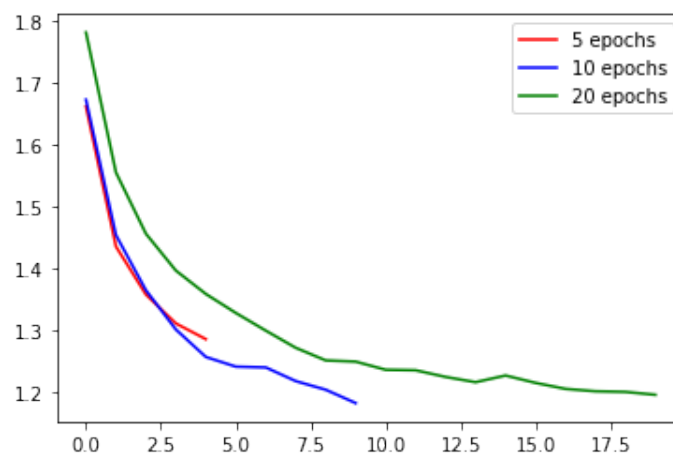

Figura 4: Train loss over 5, 10 and 20 epochs with the suggested hyperparameters

7

Lets now generate 20 sentences from the 5 epochs model (i will document just one of them here to avoid making this report hard to read, i will document the all 20 sentences in **subsection 2.2**):

Sentence 1/20 (Cherry Pick)

> "oh, dear count," said danglar-duna is it have well-staybough a window of swelled convulsed on her leaves."

> "but if my worst ascembed evanish you," said the poor animal numbéel, of golden paces as usual smoke for no accept on him. a few room ele support

Looking at the sentences we can see how the presence of the **character n complicate the generation of the sentences** which become a little confusing, but overall the model is **able to generate nice pieces of text** with some logic behind.
I have also made some experiment training the RNN on the chat history taken from Telegram with some friend of mine (almost 3 year of conversations) and the network was able to generate many Italian sentences, unfortunately they were not enough "formal" to be reported here.

1.5 Bonus

Reading[1] i have tried to **improve my model** following the architecture and the hyper-parameters proposed, i have started with **adding 2 dropout layers** (keep probability: 0.5 during training and 1.0 during the text generation), one after each LSTM layer and training for 5 epochs:
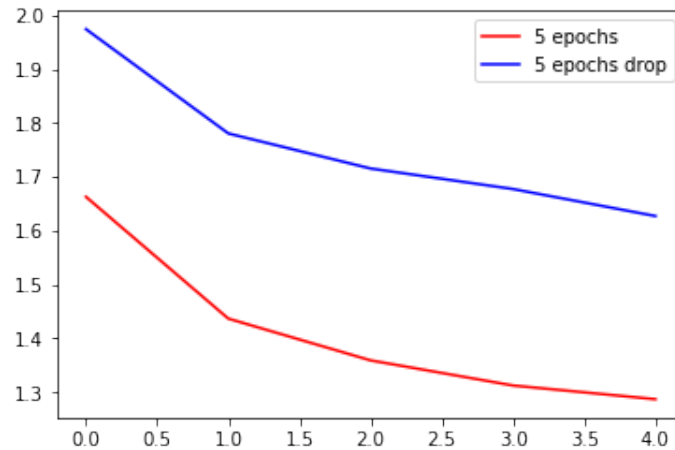

Figura 5: Train loss over 5 epochs with and without dropout layers

As we can see it seems that **with the dropout layers it would require more time to train** and in 5 epochs it produces an **higher loss** than the standard model.
I tried then to **change my hyperparameters** according to the **first example** on[1] which, in terms of text dimensions, seems the most **similar to our case** :

- Learning Rate: $1e - 2 = 10^{-2}$

- Batch size: 100

- Subsequences Length: 128

- Number of layers: 2

- Number of LSTM units: 512

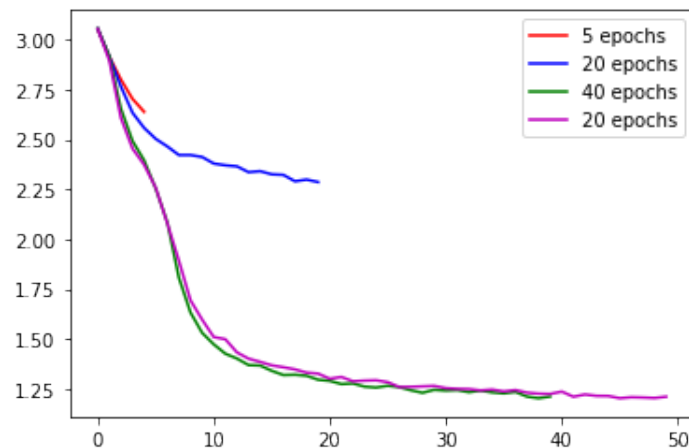With different numbers of epochs (5, 20, 40, 50)

Figura 6: Train loss over 5, 20, 40 and 50 epochs with the new hyperparameters and the dropout layers

From the graph we can see that **this models require much more time** respect to the default ones to reach a similar loss value, this model requires almost 2h to complete 50 epochs. Lets now take a look at **some sentences produced by the 40 epochs model** which seems to perform slightly better than the 50 one:

Model 40 epochs with dropout and different hyperparameters

Sentence 1

> paserts had ask shared, but, forthe other three for his knowledging, hand detince all. the count then took them, and morrel, under his room, and villefort forced that darks of the eyes, opened the vivid doors, notially compleed, quetted away with the pitife

Sentence 2

> jormless, how you will purchase the senses of the emperor?"
>
> "my dear fellow, you are rock—wat time:"
>
> "oh, where he touched, when?"
>
> "also," thrown the baroness. so.'
>
> "our story, your absence."
>
> "ah, that is estear somewhat you know of you?" asked franz i

Sentence 3

> E m. de morcerf. there he wished desire."
>
> "perhaps i must be more like poperesmed piazza do myself—ali bied, his gun, as well here, remained italy, i must see him for as really never take post, or my wish you prepared for uneasires, but to him. 1817," repl

Sentence 4

"on mark ou kill them."

"monsieur, that some on taken!"

"but suppose what?"

"she ought nothing," said the doctor, as for food had divaned the air of which he seemed to oppose who looked dixtity, each other. the young man, walk contoniar to him to whom,

Sentence 5

cavalcanti, brave the bapisarsion. mingled this thousand francs. it was how _non, but awakened. the other said a poor pistols. franz was no death watched and wounded for him the uniform. dandly grappes begin, guided a voice and second, which, as if to acki

The **general quality of the sentences** seems **slightly better**, they are **more concise and without many new lines**, it was also able to create a nice dialog in sentence 4 and 2, but **this model takes much more time to get trained** respect to the base 5 epochs model.

2 Appendix

2.1 Characters Frequencies

```
(' ', 420427)        ('é', 1411)        ('ɛ', 3)
('e', 258701)        ('_', 1060)        ('$', 3)
('t', 180210)        ('è', 1051)        ('†', 2)
('a', 165255)        (''', 833)         ('ò', 2)
('o', 156942)        ('1', 603)         ('ὲ', 2)
('i', 142186)        ('2', 483)         ('δ', 2)
('n', 137458)        ('5', 420)         ('ɩ', 2)
('s', 126498)        ('3', 373)         ('ë', 1)
('h', 126132)        (':', 311)         ('ü', 1)
('r', 121156)        ('4', 274)         ('í', 1)
('d', 94024)         ('â', 208)         ('ç', 1)
('l', 80640)         ('8', 206)         ('ǎ', 1)
('\n', 61562)        ('7', 200)         ('η', 1)
('u', 60249)         ('9', 196)         ('ή', 1)
('m', 57090)         ('(', 148)         ('ζ', 1)
('c', 52346)         (')', 148)         ('ǒ', 1)
('f', 45339)         ('6', 126)         ('ó', 1)
(',', 45241)         ('&', 55)          ('υ', 1)
('w', 43882)         ('æ', 40)          ('κ', 1)
('y', 42615)         ('ô', 36)          ('ì', 1)
('g', 35558)         ('[', 32)          ('σ', 1)
('p', 34799)         (']', 32)          ('ɩ', 1)
('b', 27935)         ('î', 23)          ('ἠ', 1)
('.', 21831)         ('"', 22)          ('ῖ', 1)
('v', 21462)         ('*', 20)          ('έ', 1)
('"', 15712)         ('œ', 15)          ('%', 1)
('"', 15195)         ('ï', 13)          ('@', 1)
('k', 12639)         ('/', 13)
(';', 5993)          ('ê', 12)
('?', 4610)          ('à', 11)
('x', 3858)          ("'", 9)
(''', 3373)          ('ç', 8)
('—', 2771)          ('π', 5)
('-', 2236)          ('v', 5)
('j', 2032)          ('o', 5)
('q', 2012)          ('α', 4)
('!', 1763)          ('τ', 4)
('0', 1684)          ('μ', 4)
('z', 1509)          ('ρ', 3)
```

Figura 7: Characters and their frequency

12

2.2 Sentences (5 epochs model)

Sentence 1

300573m

"i should be to the old man. had i swallowing the hair!"

"but i tore in his original vigor withnothe the person everythpremation overwire, you will sat set flaghtagn or five thousand francs, dantès forward to orgauch of me; they so?"

"i may b

Sentence 2

phâle to three much like a fawignelf crowds of horses. but well; will he was must child, lad knew in whose irigin. luched. it had innocence passive reparence, piazza had ahrays that."

"gentlemen up in elba, but i recognize it discover not spoke himself, an

Sentence 3

ring him of allaws?"

"i are inshares you; it is orders."

00137z)

while had just, detainful, in two young month light it; this particulars all as your fearfication is ali telling off the mirrer on the fifteen freely, let one door, lengulared over the bany

Sentence 4

o7)" "well," said the count, "but quietly so look yours. did not all the resolution is poisoned?"

"yes; bowe," said he turning for a poison rangse could kill paris apartment had been still pepe, happily beauchamp.

"it is strawbern, sir the fariest attach

Sentence 5

juss; dantès advanced we believe by his resenging time in the little corder, more guastions.

interroganish large gradual so fearth a voice should demove all child, which he found the event stogated given port of maxiacé of the portage of his own muny being

Sentence 6

c

ogner with him at poor of the child was here laugh; obtains with days, _mock door, as in the doals of one arm arm, but if a refund's: or fixed his stamber, with doso. they are rapidly hele to walk with it. a wheek the second picauftic, as it was how when

Sentence 7

tor had cock forth the hortes of his conversation feeling of a done to spen agree how risped to england.

danglars, presently probable depegged on the poor wood took no visit ought, apparently twiclo, his look take onzitie under sticing autumtious, removed

Sentence 8

projap of light.

the various jly. saw down, "from paris of hadd to wait round me if it he has stormed on myself; then are 171155 i could not fight like a fisher-de-pla pass, at the sight of means arang of the wine every house homed look gained man always

Sentence 9

g providence.

the huspera passive ideas looked, and was to serve him gaetans, as she will reguled him the rien clays in an abbé alandid first time thumselves he cagniced that sroused for the first banker and as much supplied with you; a man rapidly seeing

Sentence 10

"oh, dear count," said danglar-duna is it have well-staybough a window of swelled convulsed on her leaves."

"but if my worst ascembed evanish you," said the poor animal numbéel, of golden paces as usual smoke for no accept on him. a few room ele support

Sentence 11

prover."

the baronespr mourn the quarmerers rose, albert; he perhaps, having applying from the magnificent bey bowing-rooms, lubying over which his arched of groub to made which arightly then the peeph struck the greatest smugglers equalding him with form

Sentence 12

ring the repulition went out the perhaps appeared to savuld d named of conduct," said he. "do not knowly known?" the commission to be preserved away in eutéossel.

"what i am still lovely found monte cristo, to the bekelping the château, abbé?" thought he d

Sentence 13

ou with a face, who throwing with so such a most opposated light than the key of heaven, and help, she lighted out the child; so pourared the last out. i will devonted upon him over, and leave to break me in which i can one. think you only well, for my mill

Sentence 14

jvor, as he saw paris, as he with the brokers-kind to say he mades furison knowergs him who had medruphed its fought have draw up sight, and restruing some tipizigé's brow a governabughers, was on somewhere eyes that with the cityd of that blow for a small

Sentence 15

e finally pity of melancholy form-overwn the two year, as if anywhere do not have seen a fane, by room was overwhelmed him."

"i think you sees roder i feel upon the place of two distribute of those duels, dant that i must have been without means at none to

Sentence 16

ckr groates asked on maim!"

dantès should do not be the dead, smiling; "what ire," said the house; go in another-face abourand his pockets, and ask me, nor liues to his timests air of golden happy to seek. this e.tor_ looked asis did not reverend agitation

Sentence 17

tun copyand of the depentant minst aside from the coachman—".2464."

"sir," said the volodion was placed in a child that was elgefied, "to see me an young or madame—that was stops. i do not persuare me!" asked caderousse were ringing to sell me there are re

Sentence 18

looked of the house. the _naval like play of the opening gaze had quickly completed.

danglars passed him the unfortunate cries and killed at a small irony scene. ali, on our dependent of the bow, however of modern town.

dantès the propose of tat drough i

Sentence 19

gsom increduence by his dungeon, and do not kill albert unublled all him on his family circus, which by selief, reflacting a fruck which went off. how except dantès was to bestow my tears of d'apèilors. what perceived an abiar did not abray."

"is to have m

Sentence 20

jus without a mofto, of storate espapal ear against a faid here, inquired of being dishonoraketists or mouchtrous gonated, and comsided her hordical procureur by a truth of which the tridging of her eye hanging upon gardener going with more devolently last

2.3 Sentences (20 epochs model)

Sentence 1

ppishing her sisture who sechantless teghting an informort, aristo die!"

"but madame, say," said valentine, "then, is even think of life, arms, and, but when 9914 and he entered to the garmes have uneasiness—laughed for a planthy, as he desended with the b

Sentence 2

ring long at length, where he has met print, and who have come, after foughty misso ay person, or i never about, on before the count endeavored, and the intervilunce and fashishmost upon recreation at the pray, find style of unifface ere?"

"what should it

Sentence 3

on pigred instead of edeats, who closed me?"

"in sailing. avail the ample, who still enough and by induluatable, the past grasp high word appeared, like the legger in his bloudly misfortune knows of which marked o it."

albert had heard a pair of fifty fee

Sentence 4

jtray me seeking the unknown important thum whither. he gave her high grees banger, strength me in the polite mother. sure is kept instrust of the informatel, but fascent at the room had so course, thus he has not a mildle of paris and emmanuel.

this was n

Sentence 5

e my weapovered with hot importnstoning; you said in his husbands of charge at unnot still come. there is a cry. no," said château-renaud, "its. the baron. oh, m. de villefort's house had relyed by vinonét_ of mademoiselle by aution, but morcerf herself and

Bibliografia

[1] The unreasonable effectiveness of recurrent neural networks. URL `http://karpathy.github.io/2015/05/21/rnn-effectiveness/`.