Marco Galindo
ID: 310540013
2021, December 28th
Deep Learning Fall 2021

Homework 3

The following homework focuses on a text generation task with a recurrent neural network using a model with simple recurrent units and another one that uses LSTM units. The results of the experiments can be seen in the following sections. The homework was implemented with Tensorflow, while using Tensorboard to help with the plotting of the training process. The main resources used for the code were taken from the official documentation of Tensorflow available on their website.

1. RNN without LSTM
   1.1. Network
      1.1.1. Network architecture

The neural network is designed to predict the next letter given a previous sequence of characters. It is composed by an input layer, an RNN layer and a fully connected output layer. The first layer takes the input characters and maps them to a vector of 256 values. The following layer is composed by 512 RNN neurons; in this section these neurons are not LSTM units. The final dense layer contains 67 neurons, that correspond to the 67 different letters and characters there can be in the text.

The use of the categorical cross entropy loss serves the same function as the bits-per-character loss, by helping the network predict the sequence with the highest probability of appearance. For example, with the initial input “hello w” a good prediction for the next character would be “o”, setting up the “hello world” phrase. However, phrases like “hello where is the bathroom” also have some probability of appearance, granted it is a lower one. This loss makes the model “assign less bits” to the most probable case and “higher bits” to the most unlikely outcome.

The sequence length represents the length of the training examples used for the model. Training is carried out by making pairs of text from the input training data that feeds the network a sequence of characters and the corresponding next letter.

After trying different configurations, the final hyperparameters used for this section can be seen in Table 1.

Table 1 Hyperparameters used during training for section 1

Epochs	30
Sequence length	200 characters
RNN units	512 units
Loss	Categorical cross entropy
Batch size	64
Optimizer	Adam optimizer

1.1.2. Learning curve and training error rate

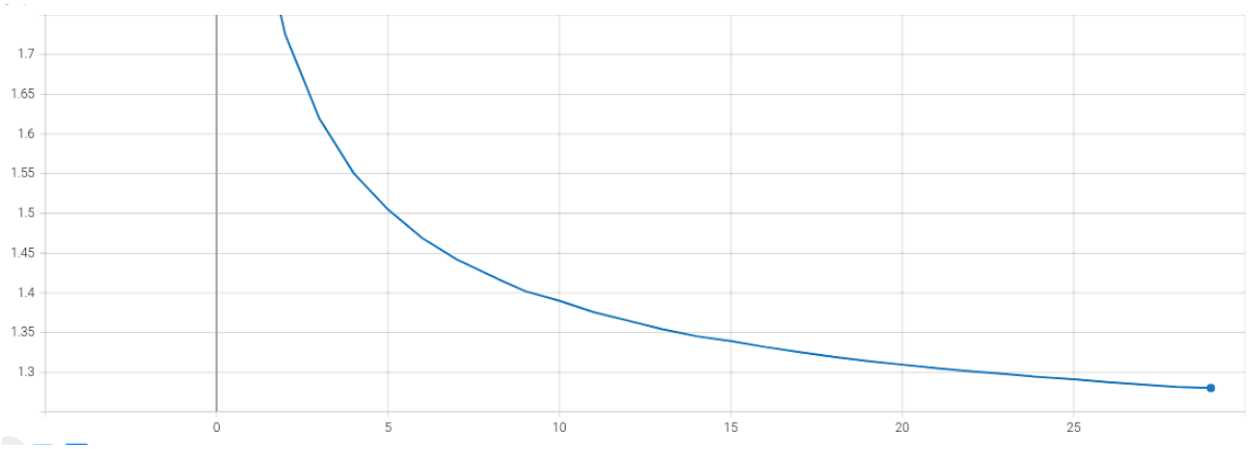

Figure 1 Training loss vs epoch for RNN without LSTM units

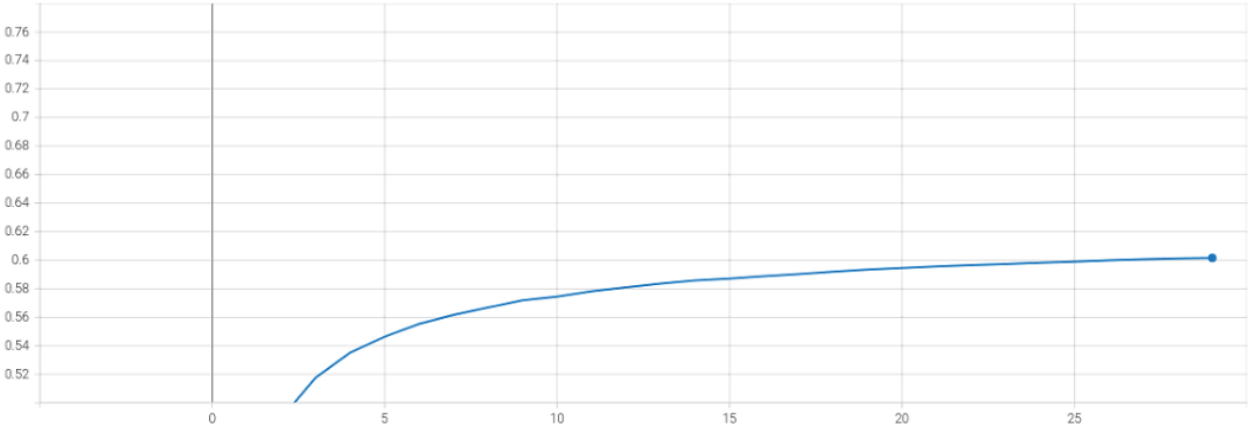

Figure 2 Accuracy vs Epoch number

1.2. Training process

To observe the progress of the model as training advanced, the new network that implemented with normal RNN units was used to generate text using the initial text "*First Citizen: Before we proceed any further*" at five different points during training. The results can be seen below.

1.2.1. Generating text after epoch: 1

First Citizen:
Before we proceed any further,
But the ood, whice her and langers.

IOWER:
Hes, we'll cemest mades armites
I stand may store
Bring grocker ead, this bear stornced-pirgaich.

AucI at your father
Haph betaly pus,
So spito much monemal.

MICK:
Her, Gies in the caint of the calls and bauth. I work's heart
Deas tike, And of I joy! decome me new in thuse
Caseet as this!

FOUD:
God yet I have weach, unwing as you, deach so I way.

SICOLEEL:
Calm, that:
Tran, I calive 'trick it: his ploverond dast, if thou thinr your paincaso ouths prrengy, ans I one fill he will and none. Gotss, Lood patcus's marnical haughter's mistserp
sit; you worment gold fross dreet,
Trurther's meancered,
Letweel we lann, on Vile; I name to have deacr's meatalk
Sometherch'd to my life his us, come, with All dos, I pay soinour feint,
Ar a Hore;
Allind the commance is tay:
And with sigh to bruded
And thise death and glain he then shase eapine enne to nemes King thyours! my lord; canestire,
This in the maid
Of thee, MyselfI's and field plains for with B

1.2.2.Generating text after epoch: 7

First Citizen:
Before we proceed any further,
How should stay' the peosomes
Shall save him to strength up.

APHELIA:
Is children much mournets shall?
Did your kindly tawn an your satury!

BIRON:
Command, a parts?
I' thir unto the formaw should be to-morrow upon
On each may paper of the witch: 'I mustress;
And to my face.

EMILIA:
What erusy, my lord, Pene; the cleir's king,
My mothouse foul king, be sits do withalt it of
Resolved, I say?

DUKE VINCENTIO:
Then, then, modeat friends,--

MINGRIUS:

Barraments, most fleeting so
much shadows thy father
Like him to make me have portune!
If my where then-gatting I hid of our arms,
And comes for matter
Enticeniwe roses,
Pardon them art thou?

MARK ANTONY:
There as shall the fools, sir?

LAURCEL:
An you so!

PAROLLES:
I help to Behelman's life: anly, madam, you make!
Strikes a vile in the carries thy bearty?

ParelyiK:
I say, I think.

CLEOPATRA:
The matter!

CALIBAN:
Never see you
To the countrabe on
most? And muldived, my causand prizes,
And yet her said!

1.2.3.Generating text after epoch: 16

First Citizen:
Before we proceed any further, or no, nay,
Now for you all; This is Captor: is 'twere friendly
As partly; the have sorrows always we'll report: 'tis good treture, thou be asquire a
Rrown's affection thou bearest it.

Nurse:
Now, by the pursued men'st, and the number naturer
That we will, and run out a Wantenolate dreams steel; doth he
which cartier with my face, in hollow.

CORIOLANUS:
You know, read.

Jowdine:
Which, Warwick, to that; for thy shown
The gentlemen From at person from your goodness,
Holds his shadow through the ladies.

CARDINAL WOLSEY:
The naw fasting to-night: you, the other. Come, and my temper to you; and he shall Reward Percy,

It would have shoption to the king.
The jewel:--O, 'sware Marcius! God favouration lamentity;
As groans that we say thy life
Should feel tweath in winges? Farewell.

First Geapted,
Encellow, yea, as the vices, either corrept these favourer that might into white
Who, before his head, swallow'd to do
In sincable ensue him
That they come not with speech
such a men of marriage

1.2.4.Generating text after epoch: 22

First Citizen:
Before we proceed any further, good counsels:
An't that caves the appleous fiend and grave-all;
And go as we lived, ere
you do see yoursel. Trop alone;
Whose does hold you
Looks dead, the world with Frenchmal and fight? Hectors. Her, that's would
have a lion's wife nearer.

Lord Chief-JOHSTONEG:
How now, Sir John I warrant her founded out
The ansista days and sisters. It is night's action
Keeps the court of Caesar?

IACHIMO:
But I protest it was led too light and by lieuping monerulous.

PRINCE HENRY:
Well, I am amazed: leave the commission.

Second Lord:
It is purse-wisdom to buy how not;
And in question good windren of the moet
That rasca say it be seen to make a gaught belongs.

SPEED:
' was with the abuned cries the deugter hath stood.
Upon my weapon, Nercation that?

SUFFOLK:
Is i' the conveniance of the foe;
And that nights here's the master;
And we'll call by giver your wife i' fantasy! will my remembrance.

OTHELLO:
I will begain you beathed hiving;
Like that my services.

KING RICHARD II:

Good gods, or law

1.2.5.Generating text after epoch: 29

First Citizen:
Before we proceed any further, if you will run
the harsh by their sinsult hived to pray.

LAUNCE:
Sir, soft, if you do marry thee,
The queen: you go off,'d thy life.
By my birds to his fair men, only shaff! and the foul heir with death.

BENEDICK:
Father, a word in hopination!
What should run here oncels he
is to render thee thee is, like an things thy claimbels,
That he has twelve dow of privately.

MACULIO:

HOLOFERNES:
He comes: who hath unto the fortunes thrive.

PRINCESS:
A half, the sea of victors thither:
By that Leona dust I heard him!

TIMON:
Trust me the little oath such as the humour of,
Not a naught to tavell'd in Alioleon: tell me
in our cock.

LAUNCE:
How he could not have made them.

BALTHAZAR:
Why I humbly mellet, betom?

PRINCESS:
A wonder. Harry, but thou shalt in an
offencisu
Thy darted retired from the pight of point. How is A merry thanks from thee!
As thought me man froth with him; and I do bulled,
Offection was the like that bred it at;
And, that breaking his eye, and cry outrull,
Let thee no

1.3. Comparison between different size of hidden states and sequence length by plotting training loss vs different parameters.

To find out the best combination of hidden state units and sequence length, different models were tested out with the same number of epochs and rest of hyperparameters. The results can be seen in Figure 3. The configurations are ordered in ascending loss values.

Name	Smoothed	Value	Step
Sequence:200, RNN Units: 512/train	1.28	1.28	29
Sequence:300, RNN Units: 512/train	1.285	1.285	29
Sequence:100, RNN Units: 512/train	1.302	1.302	29
Sequence:400, RNN Units: 512/train	1.309	1.309	29
Sequence:100, RNN Units: 2048/train	1.338	1.338	29
Sequence:50, RNN Units: 512/train	1.341	1.341	29
Sequence:100, RNN Units: 1024/train	1.344	1.344	29

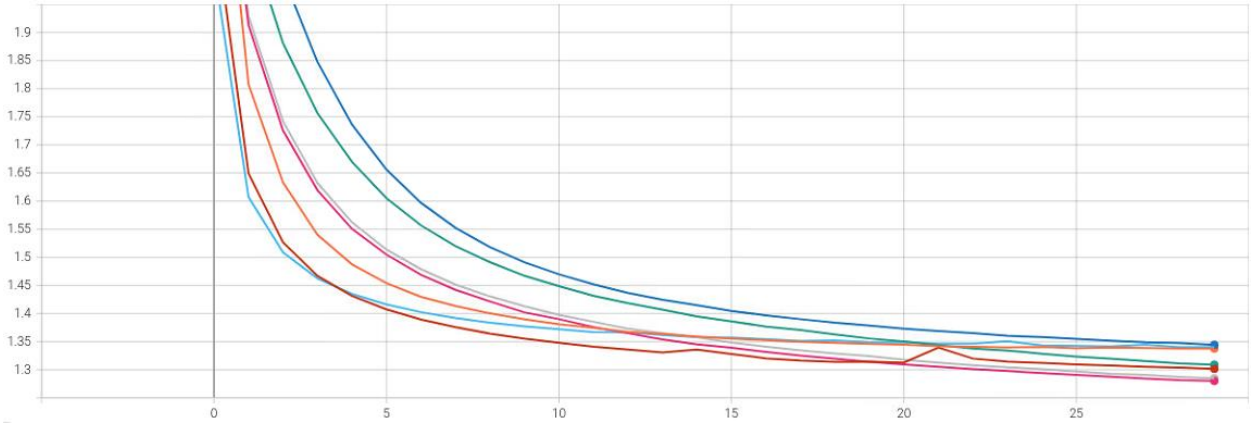

Figure 3 Different combinations of RNN units and sequence lengths

2. RNN with LSTM

2.1. Network

2.1.1. Network architecture

The neural network configuration is basically the same as the last section, with the only difference being the number of RNN units and that these units are now LSTM units.

The sequence length is also different in this section, due to the results of the experiments shown in the future section 2.3. The final hyperparameters can be seen in Table 2.

Table 2 Hyperparameters for section two

Epochs	15
Sequence length	100 characters
RNN units	2048 LSTM units
Loss	Categorical cross entropy
Batch size	64
Optimizer	Adam optimizer

2.1.2. Learning curve and training error rate

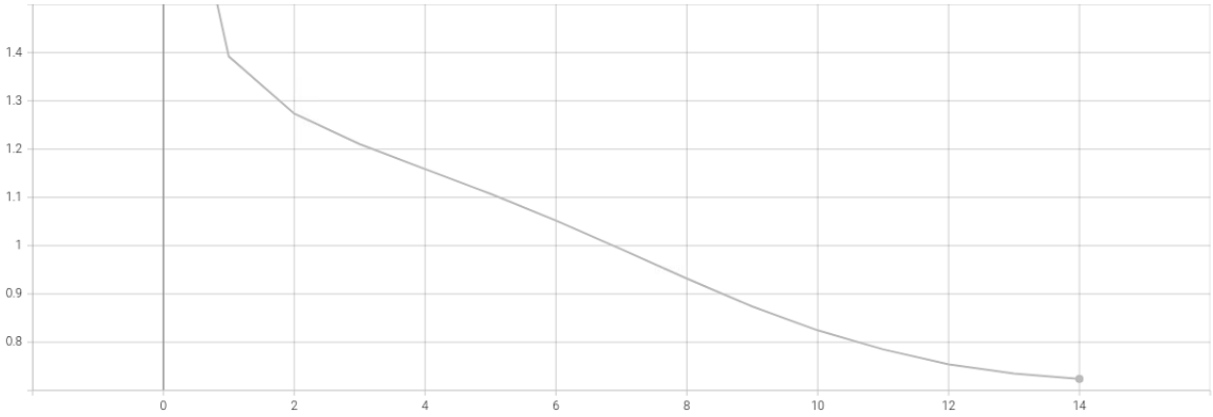

Figure 4 Training loss vs epoch for RNN with LSTM units

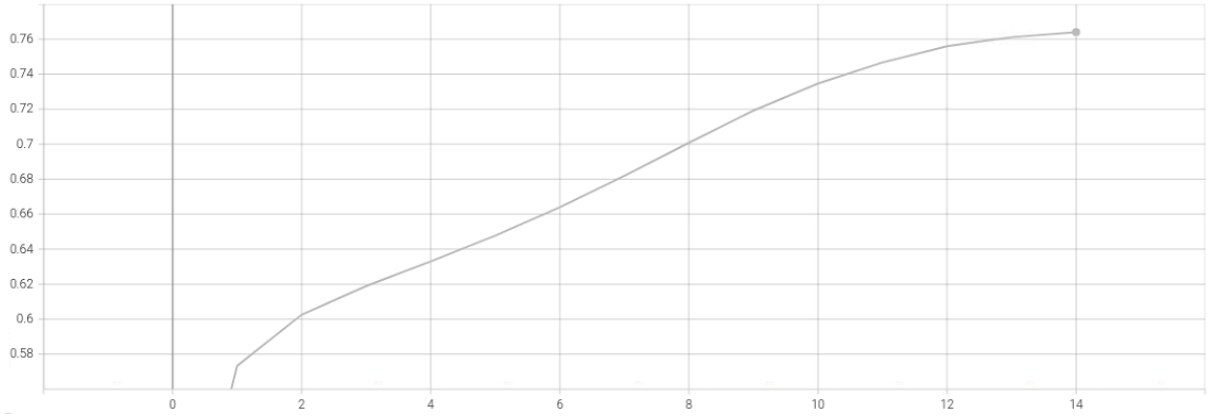

Figure 5 Accuracy vs Epoch number

2.2. Training process

To observe the progress of the model as training advanced, the new network that implemented LSTM units was used to generate text using the initial text "*First Citizen: Before we proceed any further*" at five different points during training.

It is interesting that even after just one epoch of training the network is already outputting names in all capital letters and a lot of real English words. The last remark is important as we need to understand that the network does not actually learn English grammar, it is merely trying to predict the next character in a sequence of letters. As training continues, we can see that even if the whole overall output still does not make sense, it does have long phrases that do make sense.

The results can be seen below.

2.2.1. Generating text after epoch: 1

First Citizen:
Before we proceed any further, good Falonia,
and such as the answer of you, stray, and the laging stants
When these well comes too mock, that changed as she,
In some name as
given it and deceit unto the follo?

MARCESTER:
Ay, good Master Falstaff!

ROSENCRANTZ:
Sincole his opination may his done,
Again, pretty than these fairest of to mine own,
If by best in his favours in Cambotion,
Cutsom'd his: Phesil Cofteus Lovell bad
you, marry, is it will absolu.

Second Mistress:
A, give me now mine about in tin terds! Safeity;
Strange in his old and day. Mark princess now,
And suddle redempurates, bears as that paeseth
Them so abate awhile. He speaks of thems.

Poet:
I am so thrust: thou canst be not a lief of bosilest,
and their heartily are thronged than the gain of your
arms of the commodion of the hope of Sickly,
Have rather seezs his modesty lies.'
If I call him his suitous sire to fain,
Silling aller's picture.
The dicetters, graciously the only eaten'ds
Be sweet as dobers with canquistly,

2.2.2.Generating text after epoch: 4

First Citizen:
Before we proceed any further, for I will
have younger for ever.

SHALLOW:
O God forbid it!

BIANCA:
What dost thou hear all, my lord?

Servant:
It is the soldier, that your
fat boar should fixing even in, to sing it on.

FALSTAFF:
No.

MARIANA:
Dost thou dinner trieve in thy joy?

POSTHUMUS LEONATUS:
Beleave I age, an angry hand:
He that made a fine folly gutty mock
With savage feeds rank vice, and his sisters,
Make spirit cowards than a man.

Than an unspirit of high plenty engerties
To unhatch him, and hast now gone for you.
I did not die in saying, like a case
Unto this ring: who, as a maid-butter,
To his grandfather,--loved on thee, and we make,
And carrying your patchets are a man:
Our prayers apt, last, good night: you who's himself?
Within the heavens were the enemy
To all our soldiers held-faced stabing-rest
Than all you top.

POSTHUMUS LEONATUS:
I megar, noble further, not no staff,
Nay, now forgot me.

QUEEN GERTRUDE:
Those terribles
Wilt make the dull enthralled weath and power
Sufficed with anought: hi

2.2.3.Generating text after epoch: 8

First Citizen:
Before we proceed any further, and ferrets read.
How now! what news with you?

DROMIO OF SYRACUSE:
No, thou speakest must be closed. Give her his love
indeed; for I will honour her: she was familiar to his
death, when it suffers he excuse when he beat from
me to Master Slender; the modey and your husband may
go about to overcept his argument for his
crowns concerning your more husband, jealous your preetire,
your Falstaff; by Gardon us not in this word, so
marble himself; and then 'tis not so frunty that can never
be say he knows not much.

BIRON:
A very sure: you wrong, Master Brook.

FORD:
Have you know for him? I have publed your retreat.

PRINCESS:
Screatures
And carry things graft with him. If you had but another
her, should not, whippance the profit and the tongues,
took the warrant of all Theseus, hath your grace's change is
but abhom to.

MISTRESS PAGE:
What have you made, my fair husband?

ORLANDO:
I will kill her with an unlucky man.

FALSTAFF:
Yea, as I am a Christian, then that he is too old foolery

2.2.4.Generating text after epoch: 12

First Citizen:
Before we proceed any further, Point.

HORNER:
To do what? we are good news: when your majesty
is yours, even now attempted with your worship. You are
as valiantly.

SPEED:
'Item: She brews at the gate; and look, like
angry; knock and drink; when we have
served meaning, is true.'

LORENZO:
In it and for a just prisoner, and half a dog,
and a scum lowning, which is more distrable; and the
most joined-taggemans of it passes them. If it
strips me on trialdianinks, and haste all my utteres!
hark, sir; for I am a woe is as a foolish-caterousy maid
that hath sent-folded the ingratity; that if I
shall do here again, and such as valourous as dicatemn: all
the reason that I rude my wit in the turnings of
praises: by the maid, I doubt not by adding of it;
'twas a convoyanter, or by the manner and mine:
'twill make an account we first for a man I hold
the abhorred Fife o' my trade.

QUILON:
Signior, I beseech your majesty, to desire you to dinner.

FALSTAFF:
I know him for he that hath passed all one

2.2.5.Generating text after epoch: 15

First Citizen:
Before we proceed any further, Caius Lucius
And thy friends; or the main chancellors of the
salk, is in great ere to court is immeciable: he
will make almost mons-jesters in form in counsels, as
they gave out without asks. Therefore for
France is Pagna: he is the greatest already, or a
bawd, as it were, and the best villains may prove a Christian
warrant, and howl'd in the court, thou comest not
him about the rest: they have no stomach, such as
thou hatst cost me; thou mightst be seen, of all
customs, ten too light, the king is not worth
a breakfast. Look down, rough winds and flashes!

ALICE:
Oui, vain a weak excellious dame,
No thunderAxterial, horn-in-lurious Trojan
Hath of England did commence his company.

IDEN:
Thus, old AEdrenot.
Accusat, thou candle!

MALCOLM:
I shall not.

MARCELLUMENMY:
Is thy name Duke Humphrey, in my bosom
That slew hy note? is fair?

JULIET:
Nay, let this medicine be mine, only
So gentle as a lion: thou'rt no from pompine
To the enemy, and call it perfect honour;

2.3. Comparison between different size of hidden states and sequence length by plotting training loss vs different parameters.

To find out the best combination of hidden state units and sequence length, different models were tested out with the same number of epochs and rest of hyperparameters. The results can be seen in Figure 6. The configurations are ordered in ascending loss values.

Because the model that had a sequence length of 100 characters and 2048 LSTM units seemed to decrease its performance after epoch 15, its training was halted at this point. This stoppage led to the best performance during tests, so this was the chosen final model.

Name	Smoothed	Value	Step
LSTM - Sequence:100, RNN Units: 2048, Epochs 15/train	0.7239	0.7239	14
LSTM - Sequence:100, RNN Units: 1024/train	0.8506	0.8506	29
LSTM - Sequence:100, RNN Units: 2048/train	1.014	1.014	29
LSTM - Sequence:100, RNN Units: 512/train	1.109	1.109	29
LSTM - Sequence:200, RNN Units: 512/train	1.131	1.131	29
LSTM - Sequence:50, RNN Units: 512/train	1.149	1.149	29
LSTM - Sequence:300, RNN Units: 512/train	1.177	1.177	29

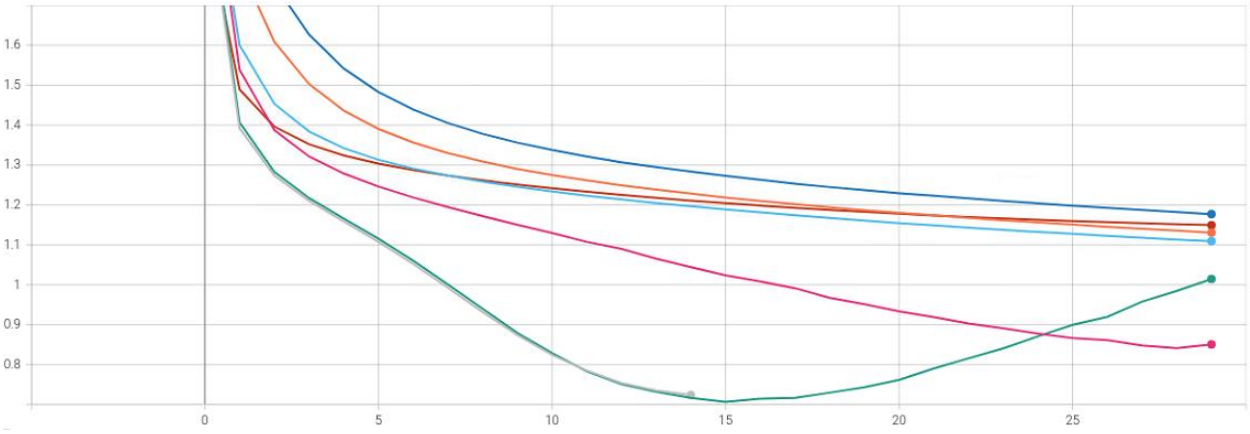

Figure 6 Different combinations of LSTM units and sequence lengths

3. Comparison between RNN with and without LSTM's

The difference between the recurrent neural networks with and without LSTM units is noticeable, with the LSTM model having better performance than the normal model. Furthermore, with the outputs from intermediate epochs, it can clearly be seen that the LSTM learns much faster than the normal model.

Another important difference is that the LSTM model benefited from a larger number of hidden units, while tests showed that the model without LSTM units performed better with a lower number of units. This could be in part because the LSTM units can decide whether to retain or forget information as they see fit with the help from two gates, in contrast with the one only gate available for the normal model.

Moreover, it is also interesting that the final output from the LSTM model does not include empty paragraphs after someone's name, and it is also able to output large paragraphs. This cannot be said for the final output from the normal model, as there is an empty dialogue, and the overall length of the paragraphs is quite small.

4. Priming the model with some input

Because the LSTM model had better performance, it was used to generate some input with "JULIET" as the starting text. The output of this paragraph is mostly in line with previous results, where there is really no connection between paragraphs, but there are still a lot of real English phrases that could be in a real story written by a human. It is also funny to see that the model immediately changed the name of JULIET to JULIETA.

Another important feat the model achieved with this result is the understanding of capital letters after some punctuation mark. After the model predicts a point or exclamation mark, it will always start the next work with a capital letter, as it should.

The result of this experiment can be seen below.

JULIETA:
If I can hear by day and night to supper.
Now, for an ush? Timorad within your sight,
Weak washed for your own desert. Let not, even for
Is past; but in the host, to death, are run.

PORTIA:
The webst and mirth o' the kingdom you shall love:
And what this goodly doctor, in Simphrook,
The Turk, that gave these two and wind, behold
Our Gaunts' estimpatives, out, ovidges down
And digg'd his thoughts how to cut off a heart?
Or rather, shall I lack a week with you;
and having that such softer after such
As when we had our kind to embark so long,
Is almost friendship: six ye hope for the time,
Whiles he may conquer fortune and reward.

EXETER:
That is most faint. As black as I have ever committed.

JULIA:
You shall be revolved if your willship that shall
beat you. Yet your cook, it was by that will she ackees.

PISTOL:
The score of court, is yet merely fought?

Second Merchant:
He did; did I nay in his pride deny
us the drink? Is there any man's heart?

ROMEO:
If he be her, now the pledges

References

- Alex Graves. (2014). Generating Sequences With Recurrent Neural Networks.
- https://colab.research.google.com/github/trekhleb/machine-learning-experiments/blob/master/experiments/text_generation_shakespeare_rnn/text_generation_shakespeare_rnn.ipynb#scrollTo=mLdnOyvzMggJ
- https://thegradient.pub/understanding-evaluation-metrics-for-language-models/
- https://stats.stackexchange.com/questions/211858/how-to-compute-bits-per-character-bpc
- https://www.tensorflow.org/text/tutorials/text_classification_rnn
- https://www.tensorflow.org/text/tutorials/text_generation
- https://www.tensorflow.org/guide/keras/rnn