LE VAN HUNG (黎文雄)

Student ID : 0860831

HOMEWORK III - DEEP LEARNING

In this homework, I will build a recurrent neural network (RNN) for character-level language model using tensorflow library. I also will try to build GRU and LSTM model to compare the results of them.

To train RNN model, I will proceed to follow these steps:
- ➢ Read data (shakespeare_train.txt for training and shakespeare_valid.txt for validation)
- ➢ Process the text: vectorize text, create targets from text, create training batches
- ➢ Build model (RNN, GRU, LSTM)
- ➢ Train the models
- ➢ Using the trained models to generate new text.

I. RNN

1. Network structure

Layer (type)	Output Shape	Param #
embedding (Embedding)	(1, None, 256)	17152
simple_rnn (SimpleRNN)	(1, None, 1024)	1311744
dense (Dense)	(1, None, 67)	68675

Total params: 1,397,571
Trainable params: 1,397,571
Non-trainable params: 0

RNN structure
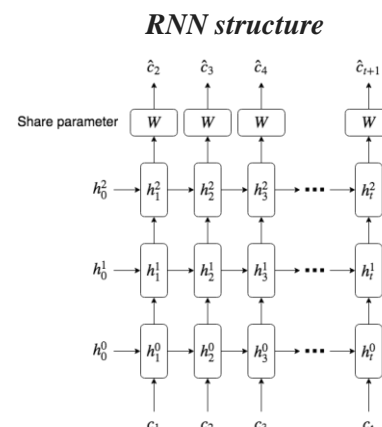
- In this homework, instead of using one-hot vector, I used an embedding layer, which is use a vector of size 256 to represent a character, which is easy to extend to a word-level.

Parameters for training			
Batch size	64	Size of sequence:	100
Number of epochs	30	Time per epochs:	~120s
Optimizer	Adam	Learning rate:	Default
Training data	43082	Validation data	673
Iterators of each epoch training	673	Iterators of each epoch val	34

2. Learning curve & Accuracy curve

Training Accuracy	60.01%	Training loss	1.29
Validation Accuracy	55.21%	Validation loss	1.52
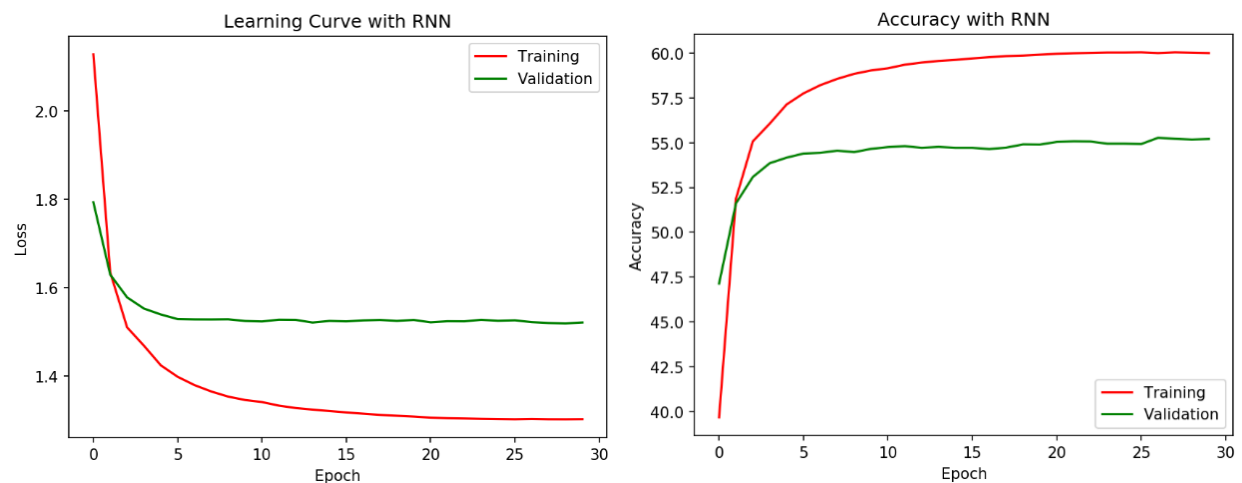
Figure1: Learning Curve of RNN - 1024 - 100

3. Validation during training

E^th	Content
0th	Accuracy epoch = 39.67 % Input : 'inity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\nof' Output : 'nity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\noff' Predict: 'ngnh oast r an elf tnd toauld te tet es tn tishter afr tf t l toyde oets tikeni tn t soatersti Tf '

1ˢᵗ	**Accuracy epoch** = 51.88 % **Input :** 'inity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\nof' **Output :** 'nity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\noff' **Predict:** 'ngnh oasdere an elf and toauld be sutned tn tish er tfr of t l thydeioues tokeny an t saatecote Tf '
10ᵗʰ	**Accuracy epoch** = 59.16 % **Input :** 'inity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\nof' **Output :** 'nity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\noff' **Predict:** "cgnh oasder' tnself,and seawld be setned tn tish ay tft of tll tuyctified hakbts\nan l seeierate Tf "
20ᵗʰ	**Accuracy epoch** = 59.97 % **Input :** 'inity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\nof' **Output :** 'nity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\noff' **Predict:** 'lgn oasder an elf ind teowld be seried tn tish ays aft of t l toyctified tikbts\nan t weaperate Tf '
29ᵗʰ	**Accuracy epoch** = 60.01 % **Input :** 'inity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\nof' **Output :** 'nity murders itself and should be buried in\nhighways out of all sanctified limit, as a desperate\noff' **Predict:** 'sgny aesder an elf,and seorld be srtied tn tish ars tfr of hll toydtified takbts\nan t seseirate Tf '

II. GRU

1. Network architecture

The setup as the same as the RNN above. The difference is I only replace layer RNN by layer GRU with the same number of units layer.

```
_____
Layer (type)              Output Shape        Param #
===============================================
embedding (Embedding)     (64, None, 256)     17152
_____
cu_dnngru (CuDNNGRU)      (64, None, 1024)    3938304
_____
dense (Dense)             (64, None, 67)      68675
===============================================
Total params: 4,024,131
Trainable params: 4,024,131
Non-trainable params: 0
_____
number of training data:   43082
```

2. Learning curve & Accuracy curve

Training Accuracy	62.18%	Training loss	1.22
Validation Accuracy	55.81%	Validation loss	1.51
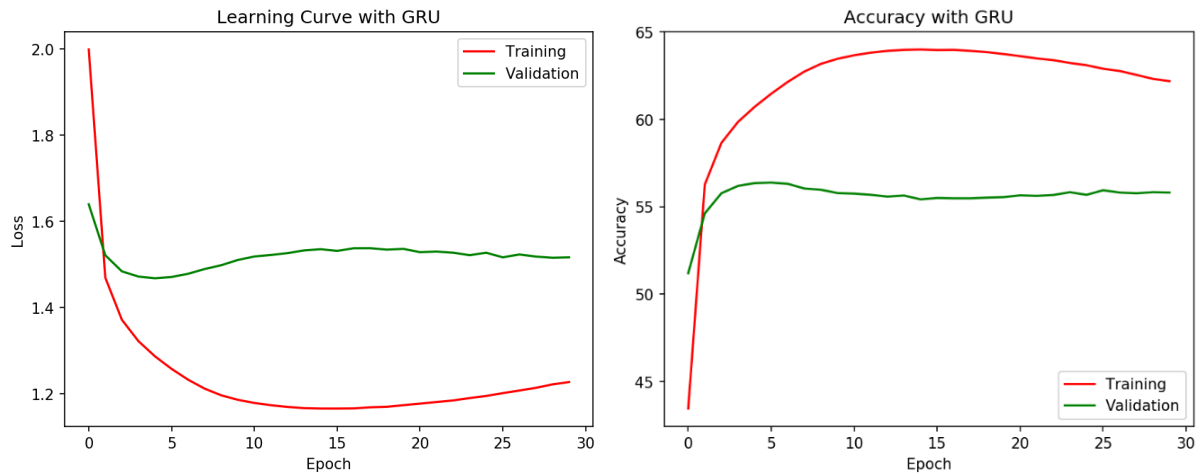
Figure2: Learning Curve of GRU - 1024 - 100

3. Validation during training

Eth	Content
0th	Accuracy epoch = 43.46 % Input : 's committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd tha' Output : ' committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd that' Predict: 'htonean ed \n\nBIRIIR:\nIot, au tynte ahet ts is \n\nBIcoon:\nWhyt ivse tollow \n\nBirst Gerch:an:\nInd thet'
1st	Accuracy epoch = 56.27 % Input : 's committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd tha' Output : ' committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd that' Predict: 'htoneitted \n\nBIRGES:\nIos, au tynt Ihet ts is \n\nBIcton:\nWhyt ivse aoll w,\n\nBirst Gitchman:\nTnd thet'
10th	Accuracy epoch = 63.66 % Input : 's committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd tha' Output : ' committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd that'

	Predict: 'hyoneetted \n\nLARGES:\nIos, au mydt thet ts ws \n\nPecton:\nWhyt ilse? sollow?\n\nSArst Satchman:\nHnd that'
20th	Accuracy epoch = 63.61 % Input : 's committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd tha' Output : ' committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd that' Predict: 'htoueet ed \n\nPARGES:\nIes, my tynta Ihet ts ws \n\nRecton:\nWhyt ilse? gollow?\n\nSArst Gitchman:\nHyd thet'
29th	Accuracy epoch = 62.18 % Input : 's committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd tha' Output : ' committed.\n\nVERGES:\nYea, by mass, that it is.\n\nSexton:\nWhat else, fellow?\n\nFirst Watchman:\nAnd that' Predict: 'hsouees ed \n\nCARGES:\nWos, mu mynt, thet Is is \n\nPIcton:\nWhyt ilse? sollow?\n\nSRrst Gatchman:\nWyd Ihet'

III.   LSTM

1. Network architecture

The setup as the same as the RNN above. The difference is I only replace layer RNN by layer LSTM with the same number of units layer.

```
_____
Layer (type)              Output Shape         Param #
===============================================
embedding (Embedding)     (64, None, 256)      17152
_____
cu_dnnlstm (CuDNNLSTM)    (64, None, 1024)     5251072
_____
dense (Dense)             (64, None, 67)       68675
===============================================
Total params: 5,336,899
Trainable params: 5,336,899
Non-trainable params: 0
_____

number of training data:   43082
number of validation data:  2198
LSTM 1024 units and len of sequence = 100
iterators of each epoch training=  673
iterators of each epoch val=      34
```

2. Learning curve & Accuracy curve

Training Accuracy	67.42%	Training loss	1.06
Validation Accuracy	55.3%	Validation loss	1.61
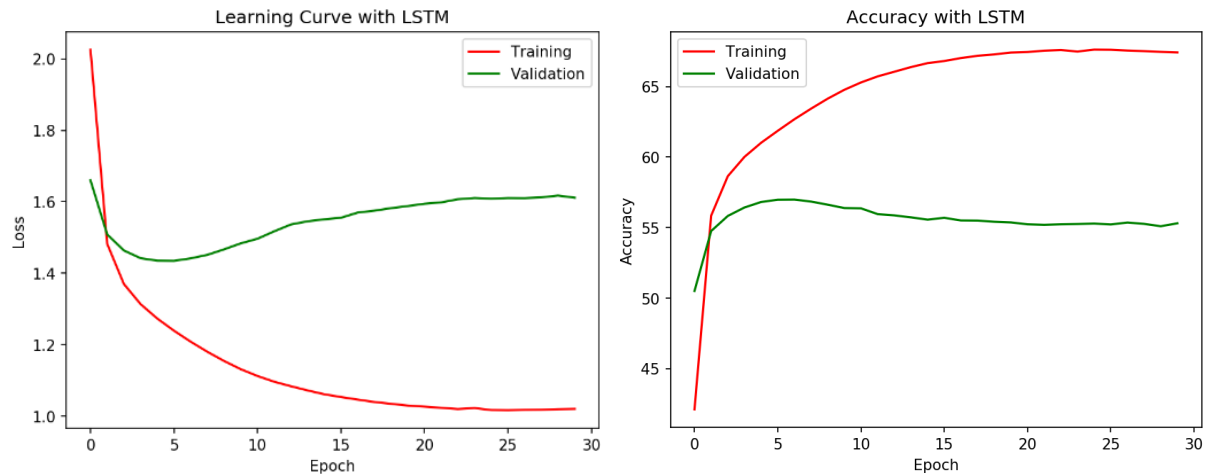
Figure3: learning Curve of LSTM - 1024 - 100

3. Validation during training

E^{th}	Content
0th	Accuracy epoch = 42.1 % Input : 'l clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\n' Output : ' clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\nN' Predict: 'dtoaeett tt.Tnmrrn oo fateur \nWesh tver e tn the e Thaane aheue thet taye the and the r sorre on T'
1st	Accuracy epoch = 55.84 % Input : 'l clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\n' Output : ' clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\nN' Predict: 'dtoam on ht.Anmrrn oo fareus \nIesh iver tn the e Thaakesaheue that tanesthe and the r sorrm os T'
10th	Accuracy epoch = 65.29 % Input : 'l clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\n' Output : ' clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\nN' Predict: 'dfoan mn sn.Tnparr oo farius Tesh aven s an sheye\nAtrike uhese that wakesthee and the r soary os T'

20th	Accuracy epoch = 67.45 % Input : 'l clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\n' Output : ' clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\nN' Predict: 'dwoam mn tt.anprir oo sasius \nTesh eren s tn these\nToruke uhese that saye them and the r clary ss T'
29th	Accuracy epoch = 67.42 % Input : 'l clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\n' Output : ' clip in it\nA pair so famous. High events as these\nStrike those that make them; and their story is\nN' Predict: 'dwoon in tt.F pair oo farius Iesh eren s ts these\nWhrikesmhise that hake the and the r paare on M'

IV. Compare RNN/GRU/LSTM with different size of sequence length and hidden state

	#hidden state	Sequence Length	#params	Accuracy Train	Accuracy Validation	Loss Train	Loss Validation
RNN	1024	100	1.397.571	60.01%	55.21%	1.29	1.52
RNN	1024	50	1.397.571	59.79%	53.18%	1.42	1.59
RNN	512	100	445.251	59.31%	54.93%	1.33	1.51
GRU	1024	100	4.022.340	62.18%	55.81%	1.22	1.51
GRU	1024	50	4.022.340	56.91%	53.27%	1.41	1.58
GRU	512	100	1.234.243	62.20%	56.85%	1.21	1.46
LSTM	1024	100	5.332.803	67.42%	55.30%	1.06	1.61
LSTM	1024	50	5.332.803	62.23%	54.34%	1.23	1.59
LSTM	512	100	1.628.483	63.08%	55.59%	1.19	1.54

From the comparision table above, we can clearly see that using LSTM layer can get the better result when compare with RNN or GRU layer. However, when we train model with lager epochs, model will be easier overfitting, so to avoid overfitting we can use earlier stopping, add penalty norm L2 to objective function J or use Drop out technique .

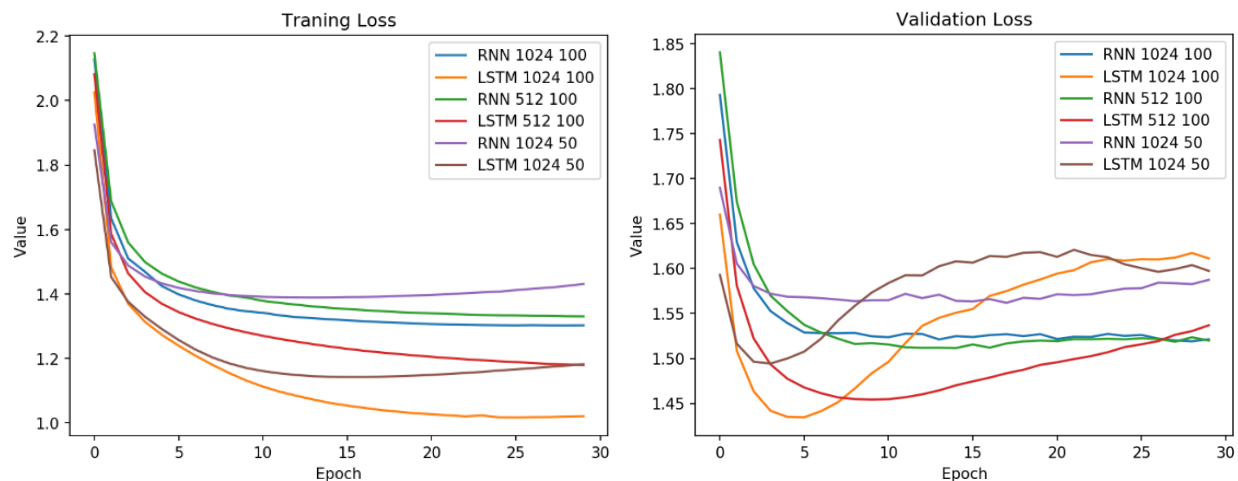➕ *You can see leaning curve of all of model above in folder: **training_curve***

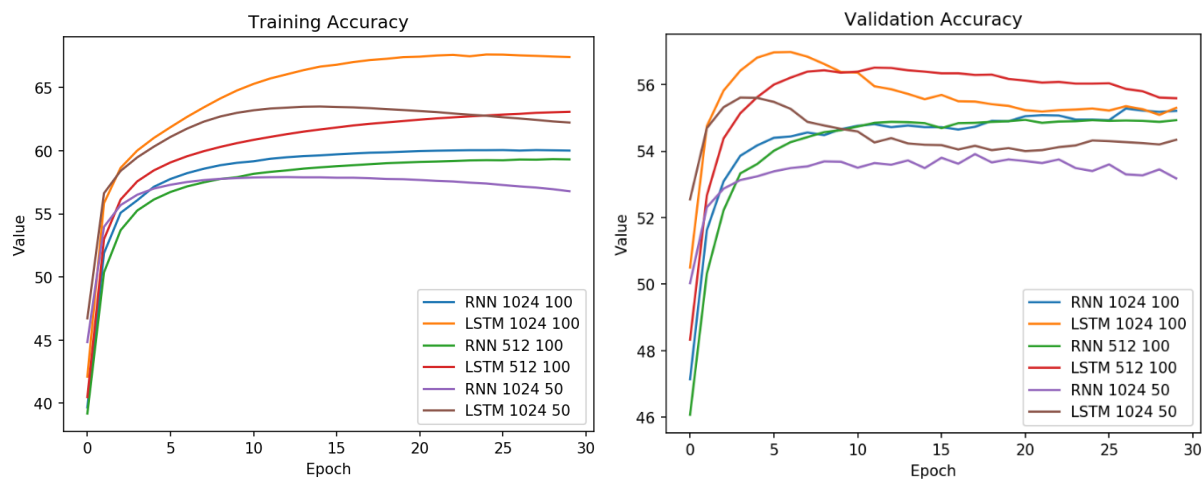*Figure4: Comparison between RNN & LSTM on loss*

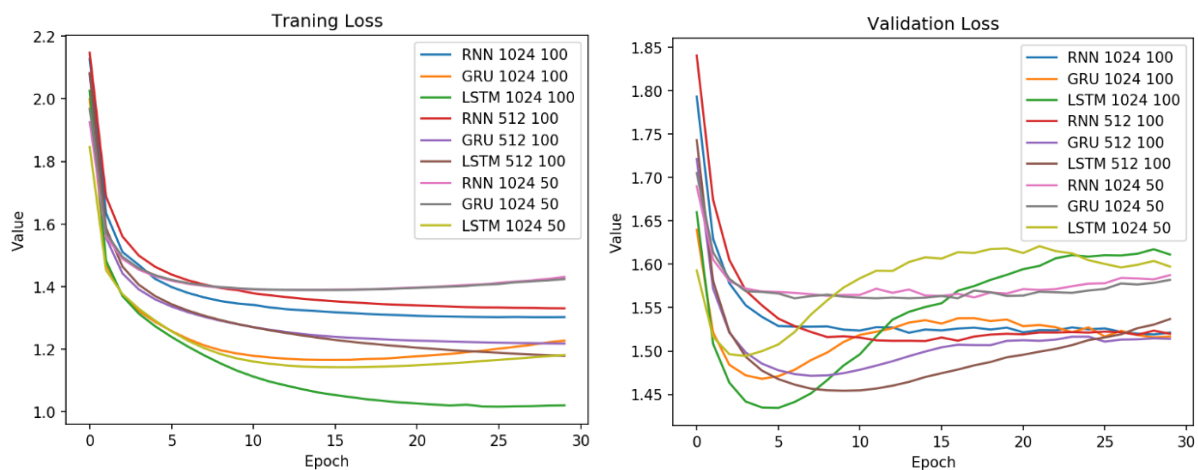
Figure5: Comparison between RNN & LSTM on accuracy

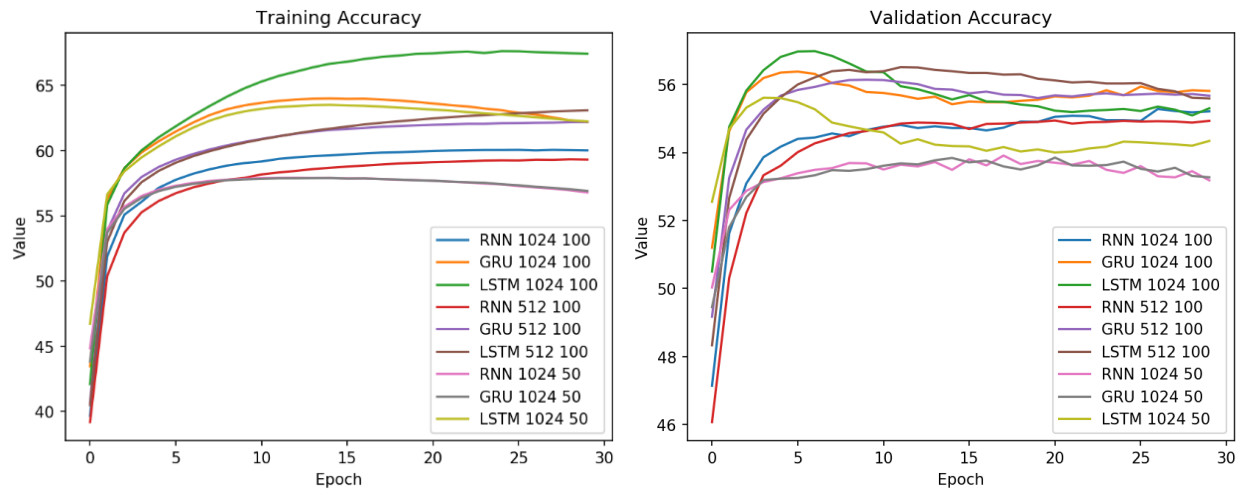*Figure6: Comparison between RNN & LSTM & GRU on loss*

Figure7: Comparison between RNN & LSTM & GRU on accuracy

V. Generate text

- I use the models with 512 units to generate text.

- Prime text to run: "ROMEO" ; "JULIET"

- Each prime text, I generate 1000 character more

1. RNN

ROMEO:
How does the hour whought,
Then they crives would find your so, or against my brother
That we of my witness to the spuler
My musical!

SICINIUS:
How ingar your friend,
Yet, that fatient, by the queen in it! he is one of
thee.

IAGO:
Wilt at leisure.

POSTHUDUSE:

I do not speak of rocks.

HAMLET:
Nay, are to bed, thwass to great grace
Where God of this dog, and they turn exambled cresking wild is hid of this dead: whom they are peace. It's o'er.

ALCTONER:
Why sends heard more than the grink to them.

Sailus.

SIR HUGH EVALICANUS:
Be not so, been said.

ANGELO:
From the poor fash; so weice which. Hang the sicer, somets in a pale course of the jures
I speak with the soft against play the
folly wounded, Hadve i' that other Lord Burgun
Wither: who Bare proadier: you terpter from peace,
Upon us on fortune, out of the general eyes,
That thou halt other
Before dead; my lawful, that I lose I did come how of Suffolk
So every cays burn of these and move ever,
Of surety smiling that Crompt'st
Th
JULIET:
None, like men did road to me exortly by Dover's days
And, did the sir!
Then is -to the great, to let's
after losse of thee: where's no
Commend him, good kne--weelia have former for false shufule's warn agant: not so:
Not our honour is he
can worn my honours, Pompeyly steep of hit: as I do aspect the patedle as he embate
Thyself: for a head, that Every shallow: shallow it in death; fair children; our orcessions
Because he should weep me half head in the most beewest unburnt us.

HAMLET:
Come inck him
upon this matter where thou mercy to me.

Offis in my ignor for't. I hope one manners shall dinner.

MONTANO:
Now, Lord Hastes you fail.

Second Gentleman:
The general victory; and your bloody ladge: the fellows this action cleth, thick, 'twould'd upon her into the ghose with my sing, as e and what;
And when you do a pock, and my heart, there turn'd the worst,
He terms off himself and
cry-honest
sing heaven. What think them. His father: if thou wretched Master Sildier, I lady: the reason

2. GRU

ROMEO:
Be falsely then and our commission:
And who is't now of Horatio deliver'd.

FLAVIUS:
This is this, I will be cheak a little man!
My haste to his mansion-filves
stirred alone.
Marry, the king! by demostis infinite same
Mischief, nor to a hair, I humbly thou hast slanders,
In unkind him, and looking gulse,
vial might hither sick, a reason is and be
Even for my honour is take to the enricly valque to anywerishments
As gives the basket in his sorrow he could
Infect Horation;
And 'twere solocted pace of cease?

IMOGEN:
Ay, villany, they shall be solecend.

TIMON:
My mistress, I say you, Desparder betin
Han, sir; he's quietning nothing; buking and a soldiers
This rhyme, with battle as if spert as vailing Tybant. What say'st thou?
But werse more: as 'tis the thoughts die.

SIMPCOF:
I have slaughted with John, in his sure I would right,
For it is wonde. Go, take my hand: cherry should I bear him,
And to live
May to think so indigal commandment of rate
And, liquidst not you in his several quar
JULIET:
He please to speak a kind of a blood.

GLEOPATRA:
Here's a word.

SIMPCOX:
What?
I say you kill this, anitch in France.

SIMPCOX:
O gentleman, let me learn that gentle Timond sons;
By the love-should,--by a power doth abume him up with her;
For suisetheavendir is so ardument.
You are justly crying on the winters?
Mohe your father E you are too.

KING HENRY V:
Brooks opinion sirecks i' the creatures
And send her leave thy conscience, mine eyes.
Sleeping continual rage day and
Were dispeased no faults and masting
Aysale aped and writ to thee, if it mean declined
So much by answer. Let me see: now have you born
to this: you follow him, on this disheman:
The enemy, both out uponation of the artian
Towards Master than I know thou wilt not assure you.

Poet:

KENT:
Niphes. I will be wholesome young MARGARET:
Why they do call? Will your has out of thy troops
For the very sea offer from thee and I
Have bound now with this fish of care, and now use France
More gives it to be put the hound,
Tel

3.   LSTM

ROMEO:
Fool,
Disdained stairs, fret devils answered within us withal.

PRINCE HENRY:
Good morrow, Might hapes.

ADRIANA:

Of natrewoman will have let no more than a' can
Are his father York should be thankfulness as the
Lioness from Roper and what, I came in us
The throat of this lapp'd Iris! you may, upon his knights colour
from thy master would have my life
His pretty devils do more for ever knit age
To give you think thou hast pass.' Monsage, do there stir
out so come to do it with feared man.

LAFEU:
I will serve, he does come against your face?
Then there
is other wither tenderness, by all
punished monicillails; the prince divary did the
violet if A you oratorous temper, then yet not see
No money of measured bawd by your chief lives? a banner,
And let it be, that friends are borne a botton to the opposed
More-faced princess better Rutlany's sake to lie.

GLOUCESTER:
It is the stake.

SILVIUS:
IMON:
He is, that my little vows, monsmil Mortimer!

First Servingman,
Thrawable fruit him on of
JULIET:
In love pursues his vow, she is not hairs in my life
to the sentence brought of a Thomas shall.

DESDEMONA:
His hands did fear me.

KING CLAUDIUS:
To whom I am hanged he moves I am; make a father have stirr'd or charber,
Whilst on the way soired and our sooner
prince do get fairest, in she plight their hands
And take it.

FALSTAFF:
Some rescued itself may think lenten off;
an all.

KING EDWARD IV: I am butter to't.

The man which he do Pistole to a blood, my flat,
Nothing. I'll turn it: with thy, for my bed with twife cry.

CAPHIS:
Come, gracious monks, his arm's nd hold return; but
put me to this turn of you; and indeed it more than a Casca: that surging yield
Proceed, having no more future,
And affoked with money: fall of suitors, hitherthe
Bucks and constable of your excellence.

LUCETTA:
Besides, my lord.

HAMLET:
Do boldetely, I pray, surely, strange,
While feeling the night waked his brow struppers, come since,
Raughters, or not you, sir.

ADRIANA:
He hath senseless battleman: if

🔢 Overfitting Problem on LSTM with 2048 unit

✓ It is quite smooth regarding to the proper spelling, grammar, punctuation.

✓ It tries to copy a phase or even a sentence from the training text

```
_____
Layer (type)                 Output Shape            Param #
==============================================================
embedding_1 (Embedding)      (1, None, 256)           17152
_____
lstm_1 (LSTM)                (1, None, 2048)        18882560
_____
dense_1 (Dense)              (1, None, 67)           137283
==============================================================
Total params: 19,036,995
Trainable params: 19,036,995
Non-trainable params: 0
```

Accuracy: 72.51%

ROMEO:
Lord John, my husband will I grant ye; and therefore tell me
what's the matter here?

PRINCE HENRY:
I know thee well.

DUKE SOLINUS:
An English lady, e they were
Some other men and means to gain themselves
As frozes in the first; for, in the presence
One, and you shall not say yourself: 'ig' is he not.
Here, Mistress Anne Page.

ANTIPHOLUS OF DAUPHIN:
We horse, Iagree, most stones, most dear;
For, being known three years old Dane;
But, not thy noble, subject'st to expect
These tedious drops.
Lord, prove plain, sola, without fair,
And budge from him that was his valiant love.

SILVIUS:
It would have spoke to-day a smiling untreadbar the likelier; slave, not d
Which you have known when I was borrowed now; but if you would have
the thing in him that will shame't strive: will you go, Marcus Brutus.

ANDROMACHARD IIN:
Yea, in both mile? will they holy arm?

ADRIANA:
Caesar doth not atteed and right through thee.

FPeet,
pray not, lessell shift nor higher than he is.

YORK:
Come, thou mine
JULIET:
No words.

IAGO:
I am sorry to you--

DIOMEDES:
No, my lord,

I'll make a monsteral fearful when I wear
Upon the very nooks to fair regards.

SUFFOLK:
What pleasant yeart there are not ignorant,
And then consort you woore on Hernk,
Would you consume and like ted so post as thou affect!
Unto the French, have never born awitch:
For what others, what does you stand halt been
That footed arm his sword with his own store
Seeming such secrets, but one of them thence
For a for a passing of it.

LUCIANA:
Will it not grapp?

OTHELLO:
Your pardon for more chains and then to glus?

YORK:
In my request deny
Thou couldst have drag ersologence is drunk. But to redemptroours in arms
And his her chamber-band of her poor breast.
Lascivers of Marchiefest the thief. Haven we looked through our
wraths, and to tave all qualifying-blowing with disposing conscience wing
Together from the former world's excell;
Were triving to his rageon, reverence,
Or sense of as a catepiris.

CRESSIDA:
How now, love's that

File for run:

File name	method
model_train.py	Train model
text_generate_best.py	Generate text
compare_model.py	Compare different structures
utils.py	File library for processing data
training_curve (folder)	Folder contain curve
data_for_compare	Accuracy and loss vector for plotting comparison