

HW3 Report - Char-level RNNs

0712238 林彥彤 YT lin

finished at 2019/12/23 14:38

1. Preprocessing of text

2. RNN

- a. Architecture, Training Record, and Checkpoint Tests
- b. Hidden State
- c. Training Sequence Length

3. LSTM

- a. Architecture, Training Record, and Checkpoint Tests
- b. Hidden State
- c. Training Sequence Length

4. Discussion

- a. RNN vs LSTM
- b. more LSTM discussion on hidden size
- c. more LSTM discussion on sequence size
- d. future works

5. Implementation Detail

- a. python version 3.6, pytorch version 1.3.1
- b. load model performs poorly(solved)

6. Appendix A: Model LSTM1024-200-30ep

7. Appendix B: Generated Texts

P.S.

I do the LSTM part first then the rnn part, so it would be better if TAs read and evaluate in the same order so that the logic is about the same timeline when I do this report.

1.Preprocessing of text

Data Loading

```
▶ data_path = "/content/gdrive/My Drive/Colab Notebooks/DL19_HW3/data/shakespeare_train.txt"
  test_data_path = "/content/gdrive/My Drive/Colab Notebooks/DL19_HW3/data/shakespeare_valid.txt"
  model_path = "/content/gdrive/My Drive/Colab Notebooks/DL19_HW3/model/"
  with io.open(data_path, 'r', encoding = 'utf8') as f:
    text = f.read()
  with io.open(test_data_path, 'r', encoding = 'utf8') as f:
    test_text = f.read()

  chars = set(text)
  tchars = set(test_text)
  char2int = {c:i for i,c in enumerate(chars)}
  int2char = dict(enumerate(chars))
  nptext = np.array([char2int[ch] for ch in text])
  npttext = np.array([char2int[ch] for ch in test_text])

  N = nptext.shape[0]
  tN = npttext.shape[0]
  Nc = len(chars)
  tNc = len(tchars)
  for ele in tchars:
    assert(ele in chars)
  print("all elememts in test data is in train data")
  print("total train data length = {}, there are {} kinds of character".format(N, Nc))
  print("total test data length = {}, there are {} kinds of character".format(tN, tNc))

□ all elememts in test data is in train data
total train data length = 4351312, there are 67 kinds of character
total test data length = 222025, there are 62 kinds of character
```

Data Preprocessing

```
▶ from typing import List, Union #python 3.8 required
  num = Union[int, float]
  def chunks(arr, chunk_len):
    for i in range(0, len(arr) - chunk_len, chunk_len):
      yield arr[i:i + chunk_len]

  def s2tensor(s: str, device: str) -> torch.LongTensor:
    return torch.Tensor([char2int[s[i]] if s[i] in char2int else 0 for i in range(len(s))]).long().to(device)

  def tensor2s(t: torch.Tensor) -> str:
    val = t.cpu().detach().numpy()
    ret = ""
    return ret.join([int2char[val[i]] for i in range(len(val))])

  #preprocessing train data:
  seqs = list(chunks(nptext, seq_len + 1))
  batches = list(chunks(seqs, batch_size))
  train_data = [torch.LongTensor(batch).transpose_(0, 1) for batch in batches]#.transpose_(0, 1) => not batch first
  #shape = (#batch, seqlen+batch, batch_size)
  print(len(train_data), train_data[0].shape)

  #preprocessing test data:
  tseqs = list(chunks(npttext, seq_len + 1))
  tbatches = list(chunks(tseqs, batch_size))
  test_data = [torch.LongTensor(batch).transpose_(0, 1) for batch in tbatches]#.transpose_(0, 1) => not batch first
  #shape = (#batch, seqlen+batch, batch_size)
  print(len(test_data), test_data[0].shape)
  #print(tensor2s(test_data[0].transpose_(0, 1)[0]))
```

2.Recurrent Network - RNN

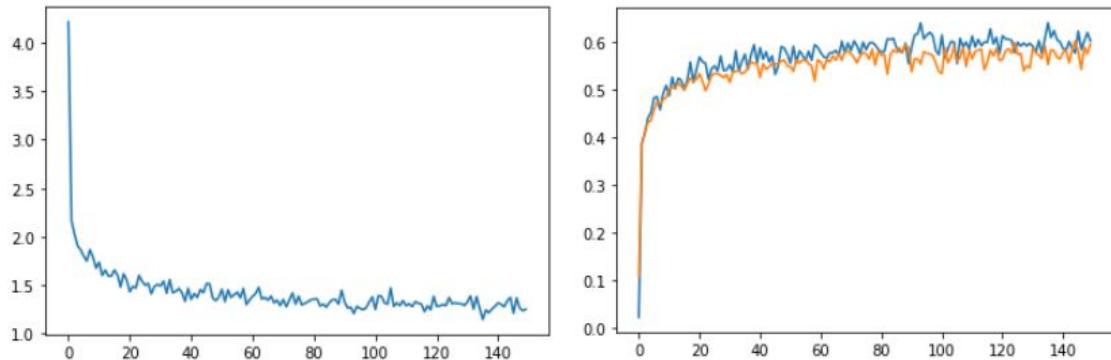
1. Construct an RNN with LSTM cell then show your

(1) network architecture

```
CharRNN(  
    (encoder): Embedding(67, 1024)  
    (rnn): RNN(1024, 1024, num_layers=2, dropout=0.2)  
    (decoder): Linear(in_features=1024, out_features=67, bias=True)  
)
```

(this is the model trained with 200-character sequence length)

(2) learning curve, (3) training error rate and (4) validation error rate



0.6014999747276306 0.5947999954223633

2. Choose 5 breakpoints during your training process to show how well your network learns through more epochs. Feed some part of your training text into RNN and show the text output.

At 1 epoch, loss = 2.435, train ac = 0.342, test ac = 0.354

```
-----input data-----  
DUKE ORSINO:  
If music be the food of love, play on;  
Give me excess of it, that, surfeiting,  
The appetite may sicken, and so die.  
That strain again! it had a dying fall:  
o, it came o'er my ear like the  
-----prediction-----  
S: I AA:  
T tystneme the sorl tf tore arai tf Aode te tneelt tf tn thet ahreert og Ahe srper oe ten thrh dt tnd th toes  
het thaann tnain  
In tev tmoeeng toll  
T tn tone tf r te lnrttoke the  
#####
```

At 3 epoch, loss = 1.634, train ac = 0.511, test ac = 0.505

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

EKE:HF EAA:

W tystcete the sarl of tove.

aray tf Aove te tncteliof tt. ahet aiceect og Aha srpeaioydoen thrk d and to tos

het thaangtnains

It iav t soeng ootl

I tn ione tf r ta lnrttoke ahe

At 12 epoch, loss = 1.380, train ac = 0.576, test ac = 0.549

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

RKE VFAEA::

T tyctc te the sirl of tove aray tf Aove me tvclcs tf tt ahat aiceeit ng The srpeaite ten benk d and to mee.

het thaang tnains tt iat n seing tols Tn tn iane tn r te snrstoke ahe

At 14 epoch, loss = 1.335, train ac = 0.585, test ac = 0.571

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

EKE:VFAEND:

I tyctc te the siol of tove

aray tf

Aove me tvclcs tf tt. ahat aiceeit ng Tha srpeaite tey benk d and to mes.

het thaang tnains It iat n seing toil

Tn tf iane tn r ty syrttike ahe

At 23 epoch, loss = 1.301, train ac = 0.599, test ac = 0.570

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

EKE VFAEADR

I yyctc ie the sirl of move, aray tn
Aove me tvcels tf mt. ahat ioceoit ng
Tha scpeaite oay benk n and to sesd

het whaang tnains It iat n seing ooil Tn tf iane tn r ty syrtaike ahe

At 29 epoch, loss = 1.319, train ac = 0.586, test ac = 0.580

At 30 epoch, loss = 1.264, train ac = 0.605, test ac = 0.575

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

EKE VFLEAIR

I tyctc ie the marl tf tove
aray tn
Aove me tvcels tf tt. ahat iiceeit ng
Tha srpaaito oay benk n tnd to fie

het whrang tnains It iat n toing tocls
Tn tf iane tn r ty syrstike ahe

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

EKE VFLEAIR

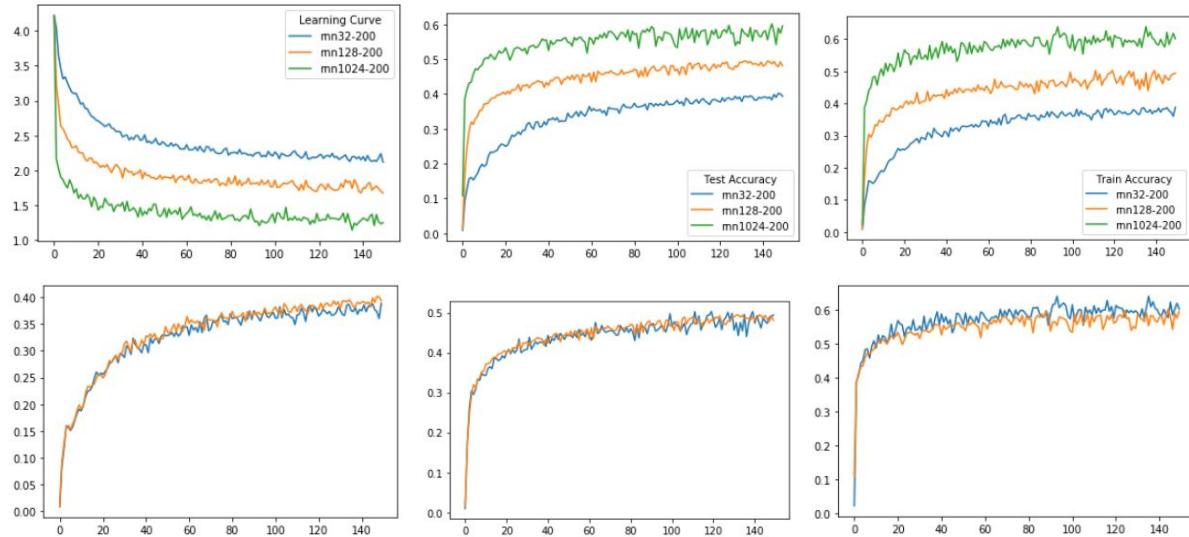
I tyctc ie the maol tf tove
aray tn
Aove me tvcels tf tt ahat iiceeit ng
Aha srpetite oay benk n and to fie

het throng tnains It iat n seing tocl
Tn tf iane tn r ty fyrstike ahi

- From 14 epochs, the testing accuracy and prediction text is about the same.
- It's because of the start of overfitting, corresponding to the validation curve at 70/140

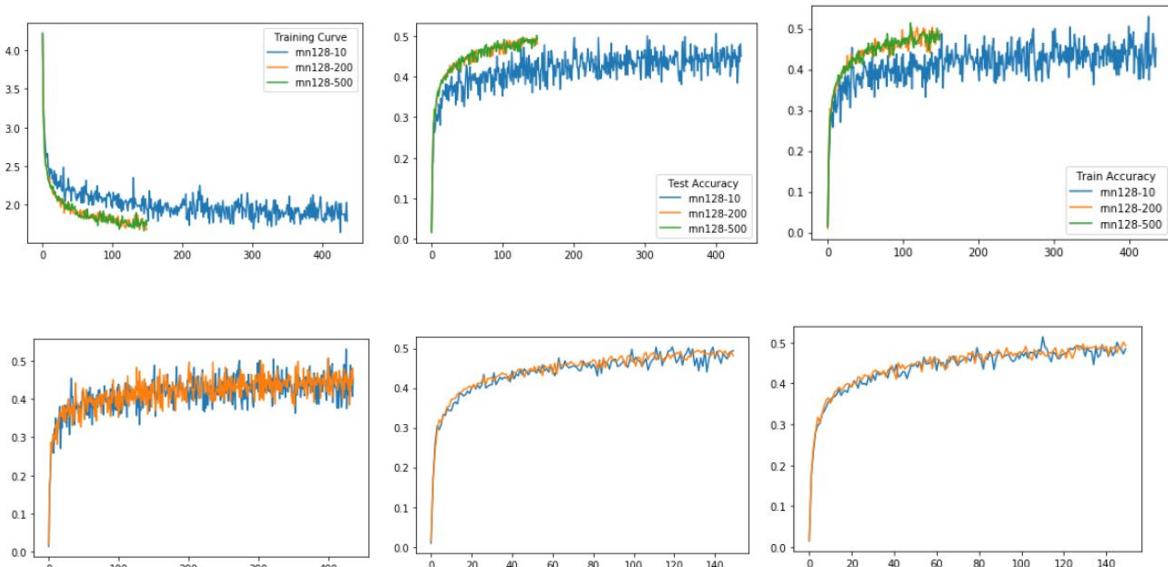
3. Compare the results of choosing different sizes of hidden states and sequence length by plotting the training loss vs. different parameters.

Hidden State:



- the second row from left to right is accuracy curves of hidden size = 32, 128, 1024
- trainning accuracy is underestimated during the process for applying dropout
- model ability increases as hidden state increases, as expected
- 32 => underfitting, 1024=> overfitting

Sequence Length:



- the second row from left to right is seq_len = 10, 200, 500
- low sequnce trainning's accuracy is eval more frequently with seqlen only 10, causing the higher variance of training curves
- sequence length can be viewed as part of model capacity for using more previous character as hidden state to predict the next character(just like n-gram model).

- So the 200/500 model performed better than seq_len = 10 one.
- We can see that when using RNN cells, seq rising from 200 to 500 change almost nothing, I think it's because of the vanishing memory problem of RNN

3.Recurrent Network - LSTM

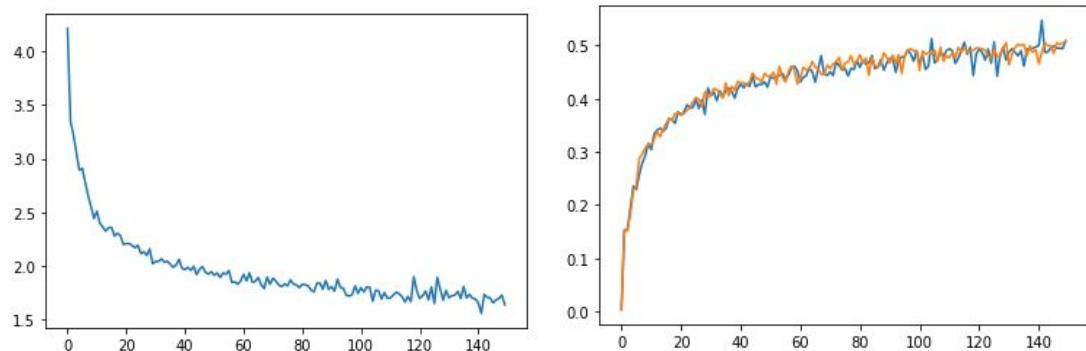
1. Construct an RNN with LSTM cell then show your

(1) network architecture

```
CharRNN(
    (encoder): Embedding(67, 128)
    (rnn): LSTM(128, 128, num_layers=2, dropout=0.2)
    (decoder): Linear(in_features=128, out_features=67, bias=True)
)
```

(this is the model trained with 200-character sequence length)

(2) learning curve, (3) training error rate and (4) validation error rate



2. Choose 5 breakpoints during your training process to show how well your network learns through more epochs. Feed some part of your training text into RNN and show the text output.

At 1 epoch, loss = 3.346, train ac = 0.148, test ac = 0.144

At 2 epoch, loss = 2.668, train ac = 0.275, test ac = 0.288

```

random test data and prediction:
-----input data-----
torch.Size([200, 1])
my maid's garments: he upon some action
Is now in durance, at Malvolio's suit,
A gentleman, and follower of my lady's.

OLIVIA:
He shall enlarge him: fetch Malvolio hither:
And yet, alas, now I remem
-----prediction-----
oo ooe tot he tee to tte
to te toe tt oot toe
oo tt to t te ooe

random test data and prediction:
-----input data-----
torch.Size([200, 1])
ee months before,
No interim, not a minute's vacancy,
Both day and night did we keep company.

DUKE ORSINO:
Here comes the countess: now heaven walks on earth.
But for thee, fellow; fellow, thy words
-----prediction-----
ter e te er
ortn n tor ntne s tenende
or ton tn tonhe ton tortor torend

AAAAEI::::::
o e tore the tortd
tor te ne tone tn trn e

ooe te tee to eeoee tee
to ht to too
on tor the tore
tore the tor t

```

At 6 epoch, loss = 2.105, train ac = 0.394, test ac = 0.400

```

-----input data-----
torch.Size([200, 1])
 pied out,
And keep it safe for our remembrance:
Return the precedent to these lords again;
That, having our fair order written down,
Both they and we, perusing o'er these notes,
May know wherefore we
-----prediction-----
ens tfr Tnd tnnd tn thle tor tfr toaaraeenge
Torhre the troaea r th the tord tnenn Thet teteng tfr torn tf r tiosher tou
Touh the tnd tir trrastng tfd the e tot r Ton know tiere ore tirt

```

At 11 epoch, loss = 1.919, train ac = 0.442, test ac = 0.444

lent us here!

QUEEN ELINOR:
He hath a trick of Coeur-de-lion's face;
The accent of his tongue affecteth him.
Do you not read some tokens of my son
In the large composition of this man?

KING JOHN:
Mi
-----prediction-----
e d tn ter

CEENN ININA::
We tavh t thoe tf tamnsssirsosn s tore
The snhontetf tes th eue tn ert rh tes

o tou wot teat th e th e tf te tomet the sovde tomeentneon tf thes tan

CING HNEO
Wys

At 24 epoch, loss = 1.744, train ac = 0.482, test ac = 0.492

At 26 epoches, loss = 1.756, train ac = 0.480, test ac = 0.490

At 30 epoches, loss = 1.681, train ac = 0.498, test ac = 0.500

-----input data-----

```
torch.Size([200, 1])
would be content,
For then I should not love thee, no, nor thou
Become thy great birth nor deserve a crown.
But thou art fair, and at thy birth, dear boy,
Nature and Fortune join'd to make thee great
```

-----prediction-----

thuld se tomsent

Aor the t whould tot tove the aot tot theu he ome the sraat teddh tot seatrve tnsoomn,

```
ut theu sre tolr and t the seddh aoar tet
Aoyere tnd trrthne tuintd th teke the toaat
```

I will not keep this form upon my head,
When there is such disorder in my wit.
O Lord! my boy, my Arthur, my fair son!
My life, my joy, my food, my all the world!
My widow-comfort, and my sorrows' cu
-----prediction-----
Tnsill tot tnep thes sor etpon ty laar
Ahan the e tn toch todneeretn ty lolh

tord,

Iy lor ty lnruese ty latr tome

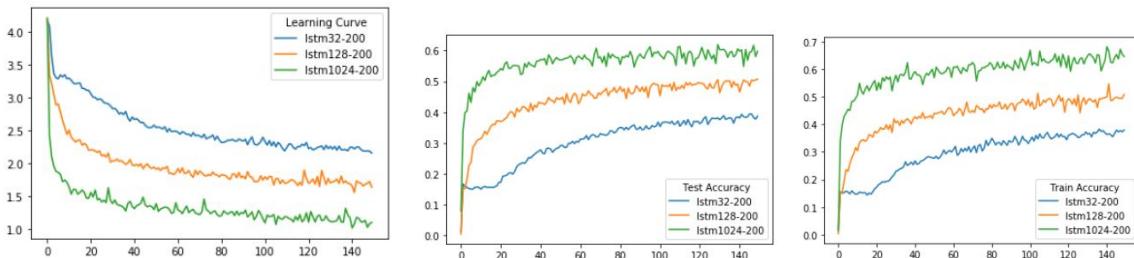
y loke ty loys ty larl ay lrl the sordd

y solen -oueort tnd ty lomeiw stor

- learn high occur rate character/vocabulary at first
- then passage structure
- then other vocabularies
- finally grammar and use of vocabularies

3. Compare the results of choosing different sizes of hidden states and sequence length by plotting the training loss vs. different parameters.

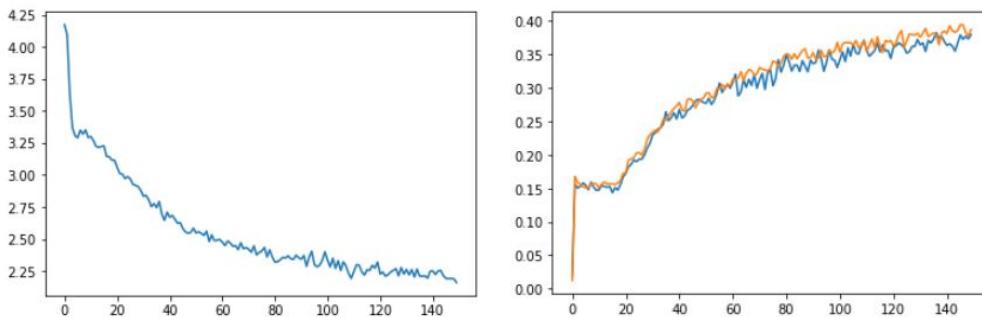
Hidden State:



- Fixed condition: trained with seq_len=200, 30 epochs each.
- we can see that the more hidden unit a model has, the more powerful it is. (without severe overfitting, can be shown below)

hidden state = 32

```
calc_accuracy(model, test_data, calc_accuracy(model, train_data,
0.3832857012748718          0.3834380805492401
```



1. model converges to CE ~ 2.2, train AC ~0.38, whole test set AC ~0.38
2. underfitting severely, the hidden state is not enough to do the task.
3. It seems that the thing it learned with its limited strength is mainly high occurrence rate vocabulary(the) and divide to passages if the sequence is long enough
4. the sudden rise around the first epoch is seen in 128/1024 as well, but then the model fails to learn more for a period of time, but at the 4th(20/(150/30)) epoch it started to learn slowly again, which is quite surprising and beyond my imagination. This phenomenon is quite confusing, hope that someone can explain why. My guess is that collab/pytorch is not working then somehow, cause the output of the test data is blank at the checkpoint 2-3.

```

print(generate(model, device, seed='juliet', tau=0.1), '\n') #deterministic
print(generate(model, device, seed='juliet', tau=2), '\n') #random

```

juliet the the

i will dorthe is enderfert if in the sather woud the thea!

CTPEVES:
Nor pintt forshen be coe?

YRIANI:
Meor hem at it us at bave, ard, I frimon, slaot. Go whade dapest thou: cuwf with for whis so, wous hath there gord hen'd me sir'sed bo'n; Bow a shen, beat pel there not he: ous to of the' do memes to for and sheam: candy kound ere, af the a there it the cich dath, Wher not to thas moud; liy hath her hall her thou and the blove sotith adferestasly beder, the my goth and dest thouw my srofe, 'hat me h

the phenomenon(epoch3(no output) to epoch 4(string of e))

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

.....Duke data.....
DUKE ORSINO
If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the
.....Prediction.....

#####

hidden state = 128

```

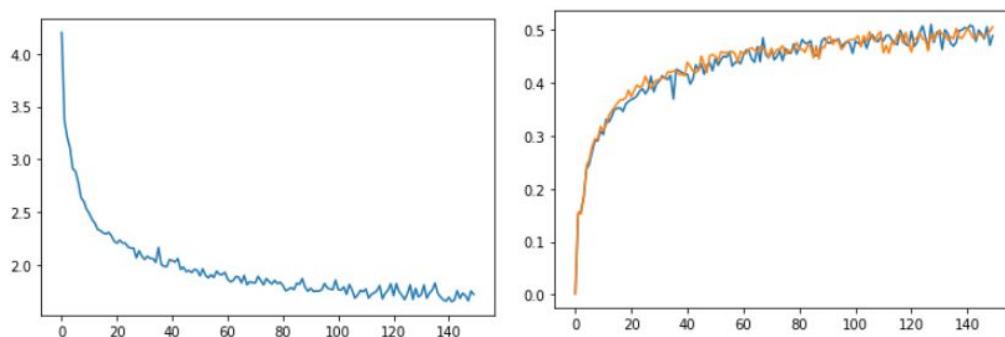
calc_accuracy(model, train_data), calc_accuracy(model, test_data),

```

0.5093761682510376

0.49467140436172485

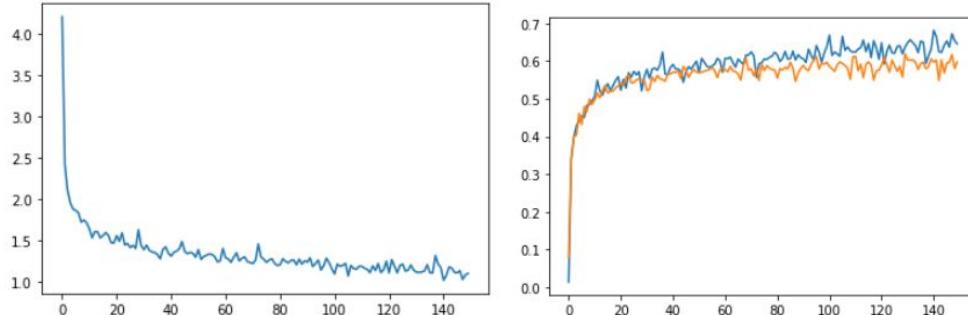
At 30 epoch, loss = 1.681, train ac = 0.498, test ac = 0.500



- Accuracy ~0.5 CE ~1.67-1.7
- It seems that 128 hidden size is kind of enough for char-level LSTM to learn some basic senses of vocabulary and grammar in the ipython notebook output.

- mode.eval() when testing, but training accuracy is calculated at the same time of training(for sake of speed), so the train and test accuracy curve are not separated by a lot. (dropout layer contribute bad accuracy when evaluating training accuracy)

hidden state = 1024



At 30 epoch, loss = 1.091, train ac = 0.652, test ac = 0.597

- Accuracy ~0.65(train), ~0.59(test), CE~1.1
- Much more training time, but significant improvement in the generated texts. It generated some context-related vocabularies and passages, which is a sign of LSTM's "long term memory".
- Sample texts are in the "HW3_LSTM2_1024t-200.ipynb" file and some in the appendix or LSTM model description.

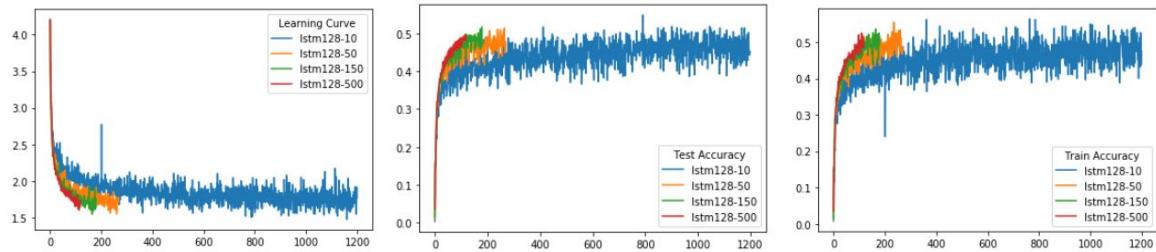
Sequence Length:

fixed conditions:

```
CharRNN(
    (encoder): Embedding(67, 128)
    (rnn): LSTM(128, 128, num_layers=2, dropout=0.2)
    (decoder): Linear(in_features=128, out_features=67, bias=True)
)
```

```
#data
seq_len = 10
batch_size = 50
#model:
hidden_size = 128
n_layers = 2
#traininng:
n_epoch = 15
lr = 1e-4
dropout = 0.2
```

cross validation

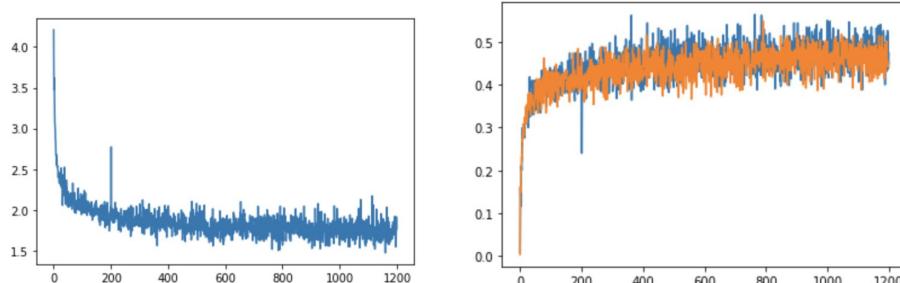


- note that because the sequence length is different, the data is chunked differently, resulting in the inconsistency of curve length in validation graph(one record per batch)
- convergence at : 10/50 - ~15 epochs, 150 - ~30 epochs, 500 - ~45 epochs

seq length = 10

(total accuracy on test data)

```
Julieter thought to me he should death and swear for himself make a mother thee,  
And be who estand the sus  
0.4732089340686798
```



```
At 15 eopoche, loss = 1.728, train ac = 0.483, test ac = 0.459  
RNN write with seed Juliet
```

```
#####
#
```

```
Julieter suck and thou be here.
```

PORTIO:

OCTAPIO:

```
And further to bent.
```

BIONCE:

```
Sir, sir, but against be of rather should he will dreps not offence the res
```

```

print(generate(model, device, seed='i will', predict_len=500, tau=0.8), '\n') #middle, produce longer

i willow the mind the joys and night the lie as return to
present to dear wrips to;
For I knows great in own an, let speak revenge and secius here in the heart was thing:
And I have
fair with bearly's sight: too reed, this guest thou strapfly saw on the commons' thee?

First Cheer.

GLOUCESTER:
Marry, if come,
That thy places it have merry,
And not thy love and isseasts, by the rest that talk, if I speak with a wars son of his stink he will not do
she doth some the virtue, Duke of craces me as wear an

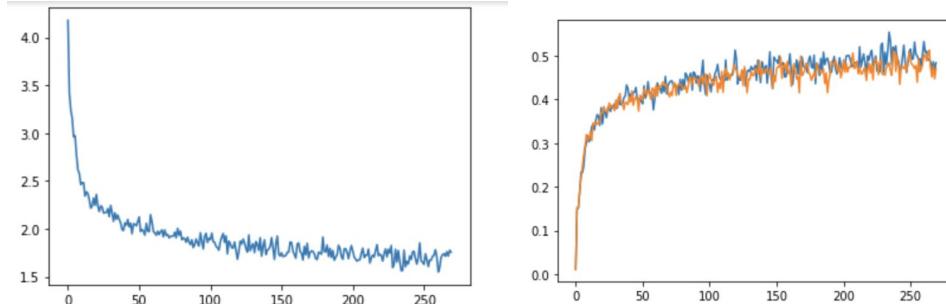
```

- higher raining time with same epoch, batch size (=>more batches)
- (At 15 ep) Stuck in lower accuracy(~0.47 on the whole test set) and higher loss(1.728)
- vocabulary: good, grammar: almost none, random passages

seq length = 50

```
calc_accuracy(model, test_data, n_batch=len(test_data)-1, device=device)
```

0.49704185128211975



At 15 epoch, loss = 1.699, train ac = 0.496, test ac = 0.478
RNN write with seed Juliet

#####

Julieter of I thank,
Which your kneed, I boy, Lessicious, through
honours he have with saidserous and bear
The tear nays, and when, who will grace.

KING RI

```

print(generate(model, device, seed='i will', predict_len=500, tau=0.8), '\n') #middle, produce longer
i willy.

TOMION:
Drave I serve my purse of my deserver of the worth
My am thow incouragement and some her leates of picks not to the purses four trouch.

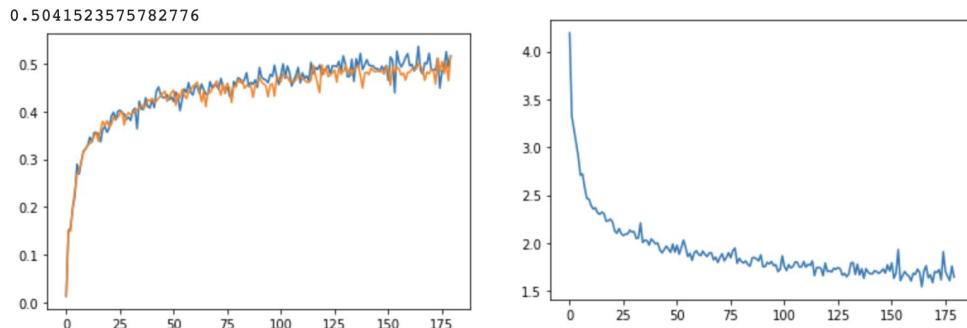
BOSTARD:
And I have hard the cuts, my lord lord,
Fissilance on of treather soth, thou with yet.

ANTIONA:
In well, he come to your kingmate,
To Jew the persawn: when he down by to his forsigly will not the
bay we comes discomel
In venty think him to have me; and taung and so lives
The liver, and it in his hibdy, that a an enement
Then the feast but hi

```

- more acceptable training time
- stop at 15 ep with slightly higher accuracy(~0.49 on the whole test set) and lower loss(1.699)
- vocabulary: good, grammar: a light sense of grammar, some passage structure

seq length = 150



```

At 15 epoches, loss = 1.713, train ac = 0.493, test ac = 0.492
RNN write with seed Juliet
#####
Julietom,
what be a lufe to all the bay the fear,
To me to hugh and pray himso sench stand?

CULEN:
All me contest of adactions.

KENT:
So mage!

COSSABI:
W
#####
print(generate(model, device, seed='I will', predict_len=500))
I will the will to the lake to mistress,
That I was to my contone one, my lord
To mound to the with to my scalles.

IATANIO:
As that Caled deselve.

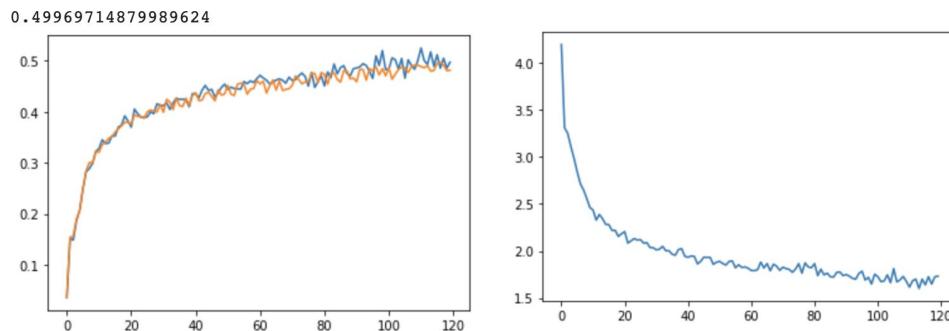
OINA:
So like, the rether way state the
charge sitton receleome! Veresount, country.

First Gentream:
Who's thee. There quees, not the world with to company to
not when to I such his basland of the ender, he
coust in the grace you wing a hisplent,
For I with but do not must be the party:
And we on the essence, sir, but the agours my more offern?
#####
MOSSTERSES

```

- much more acceptable training time, but take about 30 ep to converge
- stop at 30 ep with higher accuracy(~0.50 on the whole test set) and loss 1.66~1.7
- vocabulary: good, but some misspelling, grammar: good sense of grammar, some sentence structure, and passage structure occurs

seq length = 500



At 15 epoch, loss = 1.728, train ac = 0.491, test ac = 0.481
RNN write with seed Juliet

#####

Julietly at love excellows,
The haths in a haalter. I will that in the false bloods
Would eye of spevers to his loven to me,
Ay a and and her of it the sec

I willy wated the letters good.
For be sting sullen as the griencing our of that
ip reverent not forgerent she endied of the
Reaves of may for sleave to sand in his speak.

SROYLIO:
My lord,
For his ail father. Now so seed with so?

Hong Lord:
It send of my part to our I so mowges these to-doth
That I am sad lice of live that will see
Wiigh not and said in at mine favoices appeace
Of comphachace of demand and my lord.

KING GHARV:
Dear, I will be sup; of the weits hen
than thy heart his trival oor fit

```
print(generate(model, device, seed="I will", tau=0.4)) #Secor
```

I willed and and then I would more stand,
And she would not have me the dean to the gentle as let the
compose the name of a more to him the commons
And the father from the strumble thou shall seek,
And heart the court of the gods of the properes.

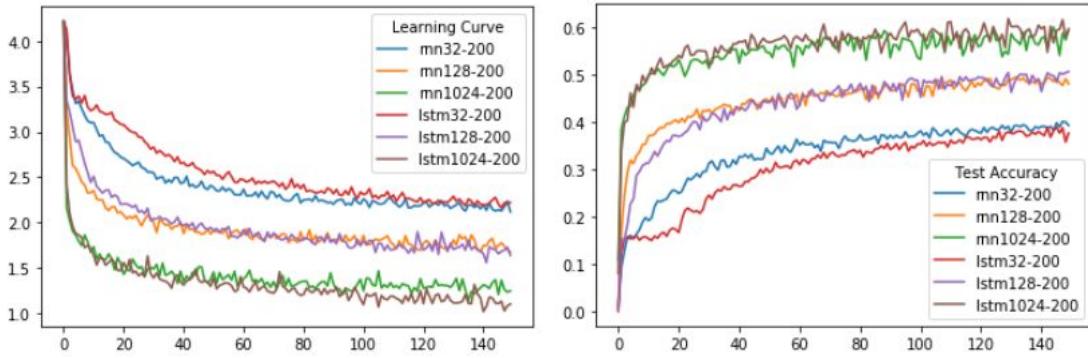
PERTIUS:
The king her for the part under the courted with the father,
With that she do the foul the best the soldier.

Second Citizen:
The and with him with a love the will as the couse
That see the did that shall de the dook:
Good more good strong and this h

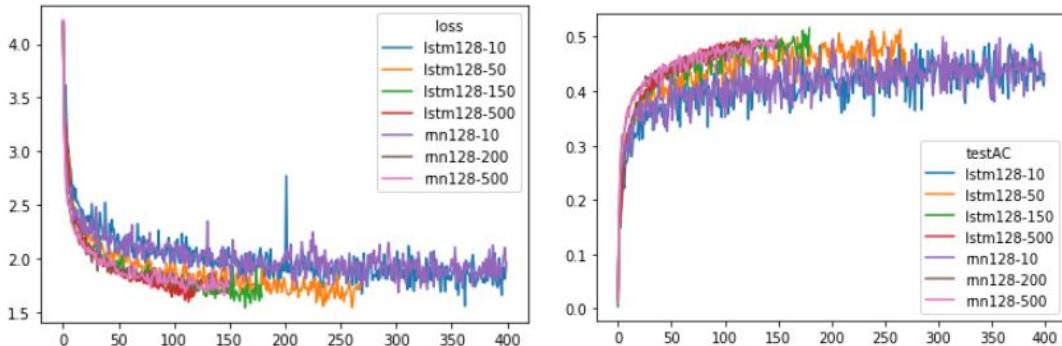
- much more acceptable training time, but take about 45 ep to converge
- stop at 45 ep with ok accuracy(~0.499 on the whole test set) and loss 1.66~1.78
- vocabulary: good, grammar: good sense of grammar, some sentence structure, and passage structure occurs, a sense of conversation. words tend to recur with distance, words around the subject with higher prob to occur in the whole passage(~500 words)

Discussion

- **RNN cell V.S. LSTM cell(hidden size):**

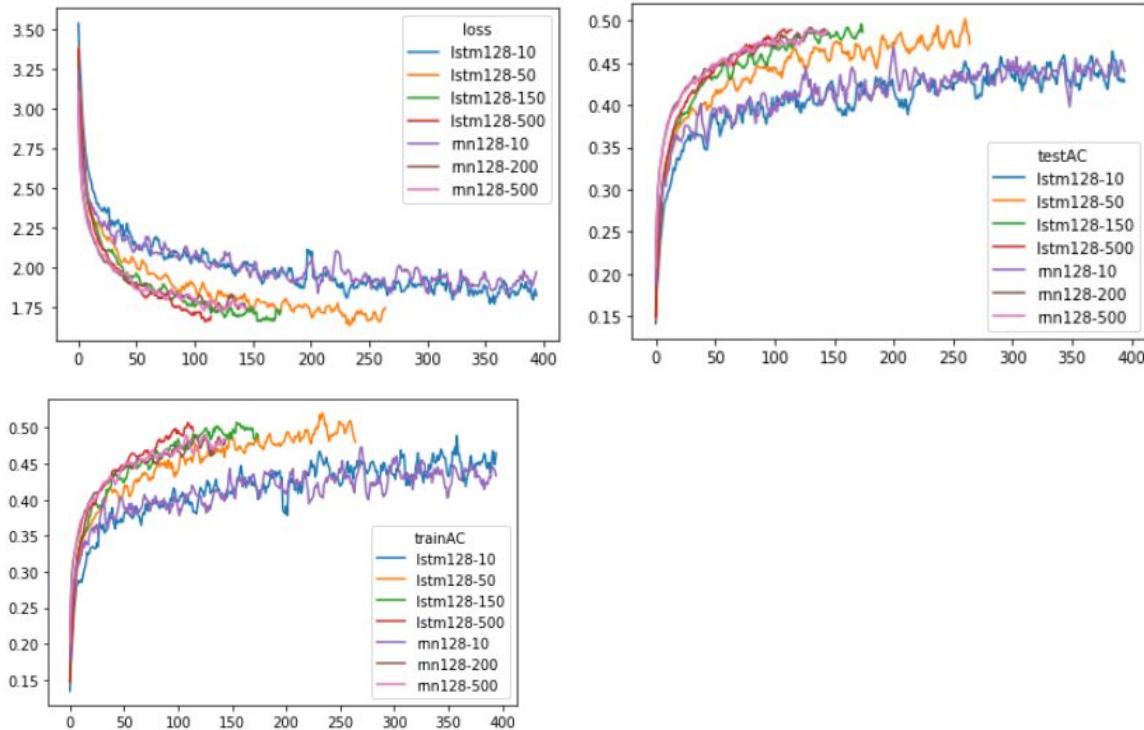


- When it comes to hidden size.
- Generally, LSTM cell performs better than RNN cell in the same conditions
- Special case : LSTM 32-200 < RNN 32-200:
 - I train the LSTM 32-200 twice, but in the 2nd-4th epoch, LSTM 32-200 stopped learning for a short period of time in both time.
 - If there is no other factors, then maybe LSTM is indeed slightly worse than traditional RNN cell when the hidden state is not enough
- Also notice that when hidden size is larger, the ability of LSTM to drop training CE further is more than that of RNN. This is by the difference in lstm/rnn1024 is bigger than that in lstm/rnn 128.
- **RNN cell V.S. LSTM cell(sequence length):**



- It is non-trivial to see the graph
- so i apply function intergrated to original log

```
def intergrated(arr, l=5):
    return [sum([arr[idx+i] for i in range(l)])/l for idx in range(len(arr)-l)]
```



- we can now see that
- Istm500 > Istm150 ~=rnn500~=rnn200 > Istm50 > Istm10~=rnn10
- Istm500 > Istm150 > Istm 50 shows that sequence length increases model capacity
- Istm 150 ~= rnn500/200 shows that the maximal memorizable hidden state for rnn of hidden state 128 maybe int the range of 150 to 200
- **Hidden State Size (based on LSTM result):**
 - the more hidden unit a model has, the more powerful it is.
 - the 1024size*2layer model start to show a behavior of overfitting, this may because char-level LSTM with large hidden size is capable of “memorizing structure and word”, but still fail to “understand” the inside meaning of the words. We can use word embedding, attention models, etc. to get better insight of Shakespeare’s work
- **Training Sequence Length (based on LSTM result):**
 - training sequence length has a similar rule to n-gram algorithm n in my opinion, deciding the number of preceding characters for prediction of the next character.
 - Training Sequence Length can be viewed as a part of model capacity
 - TA said that a model may overfit if seqlength is too long but in the case 10-500,
 - 500 perform similarly well as 150 do, and the train-test accuracy curves show no sign in overfitting.
 - So maybe the bottleneck is beyond 500, overfitting is not an issue yet
- **Future Works:**
 - test Istm 1024-500 model, which is not trained in this work and might be strong
 - try strongest model on different dataset, such as chinese lyrics, codes(even the code that generate the model, MACRO concept).
 - learn Word Embedding, Attention Model, Seq2Seq, and other advanced sequence models

Implementation Detail

- Python 3.6 + Pytorch 1.3.1
- Hyper param can be changed at beginning of a file
- Saving and Loading Model(not working unexpectedly, **solved with chars = sorted(list(set(text)))**)

Implementation Detail(Problems Encounter)

- **Should use import model.py/data.py next time** to reuse code more efficiently
- **Load model fails to work when restarting GOOGLE collab.**
- I guess that it is because of the encoding and decoding of chars is not the same, working on to fix it if I have time in the future. (the generated text with max is not repeating)
- verified by printing in the next file.....
- It seems that python doesn't have "ordered set", the backend of set may be a hash table, so every time it may have a different order, causing saved model to fail for not knowing which character is the training one. Will take about 67! to recover, or can use frequency analysis to speed up recovery of "chars" at the training stage...
- **Use sorted list from set in the future instead.**

```
modelload = torch.load(model_path+"best_model_gpu.pt", map_location=lambda storage, loc:storage.cuda(0))
modelload.eval()
print(modelload.state_dict())
```

```
OrderedDict([('encoder.weight', tensor([[ 1.0952,  0.4454,  0.6000, ...,  0.0213, -0.8545, -0.2778],
[-1.7291,  0.7577,  0.2283, ..., -0.5458,  0.9066,  0.1585],
[ 0.7658,  1.1770,  1.7650, ...,  0.6609, -1.0342,  0.6628],
...,
[-0.2598,  2.0608, -0.6937, ...,  1.0235, -1.8630, -0.5853],
[ 0.0774, -1.0873, -1.0426, ..., -0.4832, -0.3845, -1.0078],
[ 1.1430,  0.6575, -1.4067, ..., -1.1925, -0.7909,  0.5240]],
device='cuda:0')), ('rnn.weight_ih_10', tensor([[ 0.0351, -0.0130,  0.0352, ..., -0.0342,  0.0090,  0.0157],
```

```
generate(modelload, device, tau=0.1)
'Juliet&P V]C1PLrtJejxeje3ddoeBjoHendenDueFdmOnedUejexj[ej[oejexj[tnDjneYQBBehuewdehuBQuAuoeQ[enDueFdmOnedUejeYdxj[ZttL]fecPaceW$,LrtJeYQBBe[dnewuuuenDueouAQ'
```

```
print(chars)
```

```
{'d', 'e', 'R', 'Z', 'N', ':', 'g', 'B', 'C', 'Y', 'v', '3', 'c', '!'}
```

```
print(chars)
```

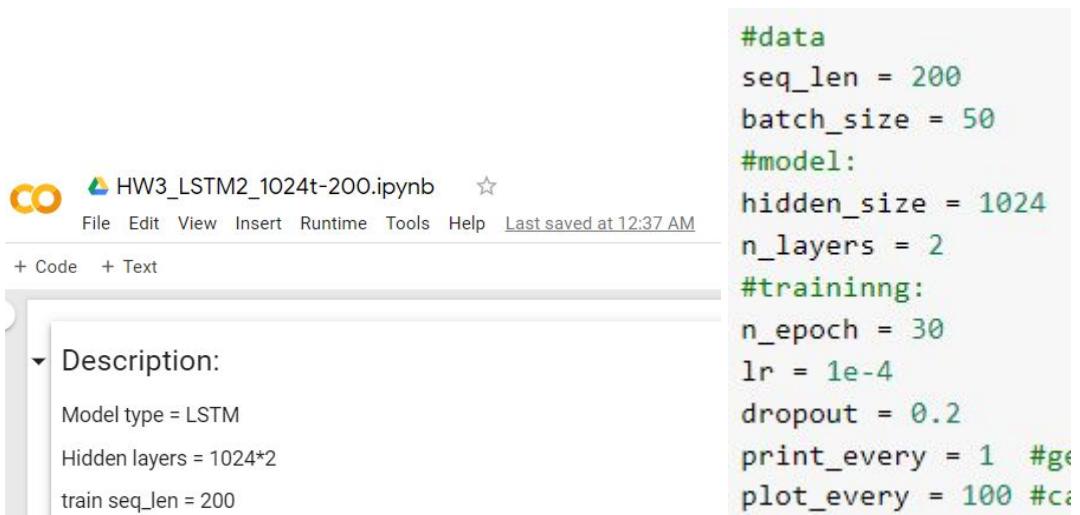
```
all elememts in test data is in train data
total train data length = 4351312, there are 67 kinds of character
total test data length = 222025, there are 62 kinds of character
{'B', 'M', 'H', 'd', 'C', 'h', 'b', 'p', 'l', 'J', 'G', ' ', 'I', 'w', 'S', ':', ';', 'D', 'X', '3', '$',
```

```
chars = sorted(list(set(text))) #set will not be the same every time, which cause loss of model before new-rnn-128-200
tchars = sorted(list(set(test_text)))
```

```
all elememts in test data is in train data
total train data length = 4351312, there are 67 kinds of character
total test data length = 222025, there are 62 kinds of character
['\n', ' ', '!', '$', '&', "", ',', '-', '.', '3', ':', ';', '?', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', 'O',
```

Appendix A: Model LSTM1024-200-30ep

- the texts are mainly from lstm 1024-200
- trainning detail (given as code)(HW3_LSTM2_1024t-200.ipynb)



```
#data
seq_len = 200
batch_size = 50
#model:
hidden_size = 1024
n_layers = 2
#traininng:
n_epoch = 30
lr = 1e-4
dropout = 0.2
print_every = 1 #ge
plot_every = 100 #ca
```

```

class CharRNN(nn.Module):
    def __init__(self, input_size, hidden_size, output_size, cell_type="rnn", n_layers=2, drop_out=0.2
                 super(CharRNN, self).__init__()
                 self.cell_type = cell_type.lower()
                 self.input_size = input_size
                 self.hidden_size = hidden_size
                 self.output_size = output_size
                 self.n_layers = n_layers

                 self.encoder = nn.Embedding(input_size, hidden_size)
                 if self.cell_type == "rnn":
                     self.rnn = nn.RNN(hidden_size, hidden_size, n_layers, dropout=dropout)
                 elif self.cell_type == "lstm":
                     self.rnn = nn.LSTM(hidden_size, hidden_size, n_layers, dropout=dropout)
                 self.decoder = nn.Linear(hidden_size, output_size)

    def forward(self, input, hidden):
        encoded = self.encoder(input)
        output, hidden = self.rnn(encoded, hidden) #print(output.shape)
        output = self.decoder(output) #print(output.shape)
        return output, hidden

    def forwarddc(self, input, hidden, device='cpu'): #for characters : torch([long]) to distribution
        input = input.view(1, 1).to(device) #batch = 1, len = 1 #need to be added:to(device)
        encoded = self.encoder(input)
        output, hidden = self.rnn(encoded, hidden)
        output = self.decoder(output)
        output = output.view(self.input_size)
        #print(output)
        return output, hidden

    def init_hidden(self, batch_size=batch_size, device='cpu'):
        if self.cell_type == "lstm":
            return (Variable(torch.zeros(self.n_layers, batch_size, self.hidden_size).to(device)),
                    Variable(torch.zeros(self.n_layers, batch_size, self.hidden_size).to(device)))
        return Variable(torch.zeros(self.n_layers, batch_size, self.hidden_size).to(device))

model = CharRNN(Nc, hidden_size, Nc, cell_type="lstm", n_layers=2)

```

```

def train():
    if not LOAD:
        optimizer = torch.optim.Adam(model.parameters(), lr=lr)
    loss_function = nn.CrossEntropyLoss()
    hidden = model.init_hidden(batch_size, device)
    try:
        #epoch_progress = tqdm.tqdm(range(1, n_epoch + 1), position=0, leave=True)
        best_ep_loss = float('inf')
        for epoch in range(1, n_epoch + 1):
            np.random.shuffle(train_data)
            batches_progress = tqdm.tqdm(train_data, position=0, leave=True)
            ep_loss = 0.0
            ep_train_ac = 0.0
            ep_test_ac = 0.0
            cnt = 0.0
            for batch, batch_tensor in enumerate(batches_progress):
                model.train() #####
                # reset gradients
                model.zero_grad()
                # data
                input_variable = Variable(batch_tensor[:-1]).to(device)
                target_variable = Variable(batch_tensor[1:].contiguous().view(-1)).to(device) #####
                # prediction
                output, _ = model(input_variable, hidden)
                output = output.view(-1,Nc) #(batch_size*seq_len, Nc)
                #loss and backward
                loss = loss_function(output, target_variable)
                loss.backward()
                optimizer.step()
                model.eval() #####
                #record
                if batch%plot_every == 0:
                    _, predicted = torch.max(output.data, 1)
                    total = predicted.size(0)
                    correct_predictions = (predicted.long() == target_variable.long()).sum()
                    train_ac = (correct_predictions*1.0/total).item()
                    test_ac = calc_accuracy(model, test_data, n_batch=1, device=device)
                    loss = loss.item()

```

- training process

100%|██████████| 432/432 [01:25<00:00, 5.09it/s, loss=1.881, test_ac=0.462, train_ac=0.440]

At 1 epoch, loss = 2.515, train ac = 0.324, test ac = 0.337
RNN write with seed Juliet

#####

Julieter,
How not so whell and to come worlqours worthou prave thee shalble live the have with at to have cause lod, proassion his lide.

LASUS:
I blow you

#####

random test data and prediction:

-----input data-----

DUKE ORSINO:
If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:

O, it came o'er my ear like the

-----prediction-----

S I :::
W tystnete the sorr tf tore aray tf Aode te ntent af tn ahet ahceern og Ahe snper oe ten thnh d and th mess
het shaann t ein In tev tmooong totl
Tf tn tone tf r ta lnrttove the

At 3 epoch, loss = 1.584, train ac = 0.523, test ac = 0.514

RNN write with seed Juliet

#####

Julieter, who when
I was you rold him be to this coldience and art good
morriel to away, but we bishinical; falled, and he was it;
And then are runch is the

#####

random test data and prediction:

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:

O, it came o'er my ear like the

-----prediction-----

S::VF AAA:

I yystc te she sarl of tove arac tf Aove me tncelsiof tt ahet aiceirt og
Ahe srprrioi oen sonk d and to sod

het Ihaang tnains It iev t meeng ootl
Tf tn ione tfer ta lnrttooke ahe

At 4 epoches, loss = 1.532, train ac = 0.541, test ac = 0.524
RNN write with seed Juliet

#####

Julietes!

HAMLET:

Now leave not to convering mad;
And to conjury he wish unto my commongue
With his thought arms are. You have her mine you.

BARMOMIUS:

So

#####

random test data and prediction:

-----input data-----

DUKE ORSINO:

If music be the food of love, play on;
Give me excess of it, that, surfeiting,
The appetite may sicken, and so die.
That strain again! it had a dying fall:
O, it came o'er my ear like the

-----prediction-----

S::VF AAE:

W yystc te she sarl of tove aray tf
Aove me tncels tf tt ahet aireert og
Tha srpeaioe oen shnk d and to ses

het Ihaang t ains tt iav b mesng toils
Tu tf ione tf r ta larsooke ahe

- almost as good as rnn1024-200-30ep

```
At 30 epoches, loss = 1.091, train ac = 0.652, test ac = 0.597
RNN write with seed Juliet
```

```
#####
```

```
Julieted and venturous lord:  
If it is not a damned my posterior,  
That we will content your over-mercy,  
Whose rags are other in the world's ear,  
And hither s
```

```
#####
```

```
random test data and prediction:
```

```
-----input data-----
```

```
DUKE ORSINO:  
If music be the food of love, play on;  
Give me excess of it, that, surfeiting,  
The appetite may sicken, and so die.  
That strain again! it had a dying fall:  
O, it came o'er my ear like the
```

```
-----prediction-----
```

```
EME VFEEAEU  
I tyctc ie the sirl of tove Irey tn Iove me txcelts tf tt. ahet iireeit ng,  
Ihe srprtite oay btnk n and to fie
```

```
het whrang onains tt ias b soing mots,  
I tf iane tner ty syrstike soim
```

• overfits train data at 20-30ep

```
random_train_seq(model, npptext, seq_len=seq_len, device=device)
```

```
-----input data-----
```

```
First Citizen:  
Before we proceed any further, hear me speak.
```

```
All:  
Speak, speak.
```

```
First Citizen:  
You are all resolved rather to die than to famish?
```

```
All:  
Resolved. resolved.
```

```
First Citizen:  
First, you
```

```
-----prediction-----
```

```
orst Sitizen:  
Ae ore te saoved i d tarther se r me speak.
```

```
SNL:  
Teak sieak.
```

```
Sirst Litizen:  
Wou are a l teaolved tether to bis fhen th siiish
```

```
SNL:  
Teolve  
Weatlive
```

```
Mirst Litizen:  
Warst, hou
```

```
random_train_seq(model, npptext, seq_len=seq_le
```

```
-----input data-----
```

```
SUFFOLK:  
That is her ransom; I deliver her;  
And those two counties I will undertake  
Your grace shall well and quietly enjoy.
```

```
REIGNIER:  
And I again, in Henry's royal name,  
As deputy unto that gracious
```

```
-----prediction-----
```

```
:RF:LK:  
Ahet ws ter fensom, a wosiver tir  
And theue tho hounsrns a will bndertake  
Tour hrace shall be l bpd tuiet y dndoy
```

```
KOGGNIER:  
Wnd t wmain tn terry s heyal fame,  
An Iiaety tpto thet sraceous
```

Appendix B: Generated Texts

At 11 epoch, loss = 1.326, train ac = 0.588, test ac = 0.573
RNN write with seed Juliet

#####

Juliet!

REGAN:

I use not so much of his bonds against.

BULY:

No an all, to my master.

CALEBAR:

So some least amission and short-like mind

Lordo, and be your brother.

IAGO:

I am not worthy than the father of the world.

DOMITIUS ENOBARBUS:

How now, my lord! what says the world is this?

Second Lord:

Well said, sir, the devil has a noble truth.

PRINCE HENRY:

Farewell, good father.

MENAS:

I am a son of them.

MARK ANTONY:

He is a good discontented father to his house,
And he will serve the world in his honour than
I have promised to be advised to him.

LADY MACBETH:

What is the matter?

PAROLLES:

And so shall I say he will desire to d

```
print(generate(model, device, seed='Flower', predict_len=500, tau=0.4), '\n')
```

Flowery, and be not as you are.

PARIS:

This is the name of Capulet, with all the land,
The heavens that were the sun of the king's face,
The other makes me not and dispatch.

DUKE VINCENTIO:

I am a good trumpet, and your son is so,
And I will do thee partisans to the court.

First Lord:

I am a fool, sir, but that I am sure
That is not gone to see your hands.

Second Lord:

I cannot tell them that you love me.

KING HENRY VIII:

And that we will be soldiers to his face,
And the commons are my father and

```
print(generate(model, device, seed='Soldier', predict_len=500, tau=0.4), '\n')
```

Soldiery:

He is a good prince, and the more than mine eyes
Be sure of honour and my brother's house;
And then the soldiers shall his face is strange.

CARDINAL WOLSEY:

The sin of money were in their own beams,
And will not speak to him that will not be a
devil.

HAMLET:

Ay, and the third of the sea-sick lieutenant.

CALIBAN:

Ay, marry, well; I say, and thou art mine.

PETRUCHIO:

My lord, you are most near the man that you
Will do as well as the sense and my grace
To say the last sorrow is a man of him.

```
print(generate(model2, device, seed='My lord, ', predict_len=500, tau=0.4), '\n')
```

My lord, the prince hath made him seem
To let him give his hand to his descent.

SILVIUS:

She hath been good morrow, and the state of men
Are not so near the chain that is mine own.

DUKE VINCENTIO:

What says the time?

DOMITIUS ENOBARBUS:

He shall not be the warrant of his son.

SUFFOLK:

Most fair and profession, if you have me speak,
But startly as he was a great summer,
And thou shalt know me in the court here in
As well as I am here.

ALL:

And here, my lord.

CASSIUS:

And what will you do it? I have

My lord, the prince hath made him seem
To let him give his hand to his descent.

SILVIUS:

She hath been good morrow, and the state of men
Are not so near the chain that is mine own.

DUKE VINCENTIO:

What says the time?

DOMITIUS ENOBARBUS:

He shall not be the warrant of his son.

SUFFOLK:

Most fair and profession, if you have me speak,
But startly as he was a great summer,
And thou shalt know me in the court here in
As well as I am here.

ALL:

And here, my lord.

CASSIUS:

And what will you do it? I have

i willingly in the company of France, the state
in monstrous expliations and in the son upon a woman,
and his own discourse.

How! what mean you? what thinks you had the use of
the knave may be compassion of?

LEONATO:

Come, in the discretion of the bold is your
with a blank, but not a man for me and your majesty
is well as a brave point of many a thousand man
shall be counted confessions to speak a wither's
about, and there was in the blessed linen.

CALIBAN:

Ay, burn; I might see him but a house to h

i willingly die a ship behind our mother's
graces, a charm that the poor pretty complexions
of as home seems as when the waste can appear it.

ANGELO:

How dost thou never heard him?

ALEXANDER:

Here is more knave, and all the galleys of the
lady go to die.

Lord Chief-Justice:

I have a more precious to be well in pity;
there is my death from his change: you might have
great power too long with great and storms, when
his son cowards is dead and way of life. An I
will never heel through the proud minist

i willy be the streets, and she shall be
The speedy of my soul, and there is nothing
That you will stay to say so much upon you:
My lord cannot be the strong with this form,
When I shall have a sudden shame to me.

EARL OF WORCESTER:

I am a soldier, that I may speak the world
That I may see a thousand ducats that
Which should be strong and from the field of strong,
And they did see the armour of the stream
And stretch the meanest poison of the world,
That they may prove a soldier that the sense
Which

i willy prove a truth of intellegue of her leek.
I hope
Goes to the prince apprehensions all.

IAGO:

So shall you speak to do,
And yet I know the mash were punish'd for
To spite and by the false white bed-fancing
The head of birds and misery.

BASTARD OF ORLEANS:

I thank thee hence to do.

GREMIO:

Art thou their mother there?

First Musician:

Well, they are looked but that the flood is made,
How ill our stocks be something like the spirit,
We will it not my kisse in that which majested
The speech part

Lordo, and be your brother.

IAGO:

I am not worthy than the father of the world.

DOMITIUS ENOBARBUS:

How now, my lord! what says the world is this?

Second Lord:

Well said, sir, the devil has a noble truth.

PRINCE HENRY:

Farewell, good father.

MENAS:

I am a son of them.

MARK ANTONY:

He is a good discontented father to his house,
And he will serve the world in his honour than
I have promised to be advised to him.

LADY MACBETH:

What is the matter?

PAROLLES:

And so shall I say he will desire to d

```
print(generate(model, device, seed='Lord', predict_len=500, tau=0.4),
```

Lordo, my lord, have I done thine of this discourse
To see my son of this deserving man?

PRINCE HENRY:

What is the matter?

EDMUND:

He is a devil, and hasty than a place.

KING HENRY VI:

I do not like the sense of men that she
Her professes will be buried in the world
Which is too far for me to share a son.

LYSANDER:

I am sorry that I am a fool to seek him.

CORNWALL:

I am a good resolution and my sister,
And all the world is much for one in his discourse,
And therefore I will prove a friend to hi

```
print(generate(model, device, seed='Flower', predict_len=500, tau=0.4), '\n')
```

Flowery, and be not as you are.

PARIS:

This is the name of Capulet, with all the land,
The heavens that were the sun of the king's face,
The other makes me not and dispatch.

DUKE VINCENTIO:

I am a good trumpet, and your son is so,
And I will do thee partisans to the court.

First Lord:

I am a fool, sir, but that I am sure
That is not gone to see your hands.

Second Lord:

I cannot tell them that you love me.

KING HENRY VIII:

And that we will be soldiers to his face,
And the commons are my father and

```
print(generate(model, device, seed='Soldier', predict_len=500, tau=0.4),
```

Soldiery:

He is a good prince, and the more than mine eyes
Be sure of honour and my brother's house;
And then the soldiers shall his face is strange.

CARDINAL WOLSEY:

The sin of money were in their own beams,
And will not speak to him that will not be a
devil.

HAMLET:

Ay, and the third of the sea-sick lieutenant.

CALIBAN:

Ay, marry, well; I say, and thou art mine.

PETRUCHIO:

My lord, you are most near the man that you
Will do as well as the sense and my grace
To say the last sorrow is a man of him.

Soldiery:

This is a reverend care of the duke's death.

LADY ANNE:

The cardinal's doth remember the sons
And by the strength of the hour straight shall be,
And let him find him with a first heart well.

ROSALIND:

Out, that the saints of this command the man!

CASSIUS:

And what shall I see the remedy?

SHYLOCK:

I have done my lord and great and honour,
I will not die in my strange bondman.

HELENA:

How now, my lord!

KING RICHARD III:

Why, this is the matter: he is a good father,
I thank your lordship a

```
print(generate(model, device, seed='KING RICHARD III:', predict_len=1000, tau=0.4),
```

KING RICHARD III:

All that the world is dead, so far a son,
The heavens make them be acquainted with the place.

BRUTUS:

The first and white regard of your mouth
I shall be drunk and blood that show me face,
And she is not so far as you are so
As you are sent as dear.

CAPULET:

Because you are mine own dear father's house,
And therefore let me see a stranger than
What I will find him something be a soul
That I will do thee service and my sword,
But did my soul and true and play the same:
And therefore I'll sat in a sea of thine,
And there is more than a bark again.

SIR HUGH EVANS:

Good morrow, good morrow, and the duke hath brought you.

DUKE VINCENTIO:

No, not a monster.

IAGO:

What, have you gone to deny your hearts?

ANTIPHOLUS OF EPHESUS:

The cause is gone, and he was found to see him:
But how did you see how the proper spirits
Is like a dear of all the world is coming?

KING HENRY VIII:

And that will I report the story of the world,
That I should see the best of my brother,
And the devil shall be

```
print(generate(model, device, seed='princess', predict_len=1000, tau=0.4)
```

princess of my heart:
The one be sorry that I am a good
That would be suit to her and my sword shows.

GLOUCESTER:
But this is the true death of my action,
And that the state of mine has died to thee,
And find the harmless garments of the greatest
Cannot help thee from the hand of men.

TITUS ANDRONICUS:
Sir, have you been this to me and this?

HAMLET:
There is no matter, my lord, if you are so good,
And will approve me to a proper man.

DUKE OF AUMERLE:
I know not what you have done before my spirits.

TIMON:
What is the matter, what is the matter,
That this may grow to the common proper of
And the one more better than the sharper?

ANTIPHOLUS OF EPHESUS:
Then let me see your lordship and the world
I shall discover the truth of your distress.

Second Lord:
I do beseech you, good Lord Timon, he hath
companion to the earth.

SIR HUGH EVANS:
How now, my lord! what say you this? where is the
marriage of the court?

PRINCE HENRY:
What must be so? when she was to discover him
and part of his house

```
print(generate(model, device, seed='Lin Yang Tong', predict_len=1000, tau=0.4)
```

Lin Yang Tonger Harry of his heart,
And bid him still his blood at home, being one.

DEMETRIUS:

The sixth of the earth doth seek the dead that did
As he are found she lived; and therefore did
A proper shepherd show the sides of them.

HELENA:

You are a soldier, and your strength and your
Are perfection to a little shall.

DIANA:

I will speak my lord and thing that I might have done.

BRUTUS:

And that shall be a true deceit of my bond;
And there I will not live to see your grace
And the good deed are the most proud of mine.

KING HENRY VI:

And therefore let us see the cause of breath.

MARK ANTONY:

The service of the spirit of the time
Is much to stand and see the sister of the night.

Second Lord:

I did not think the devil advise him there.

WARWICK:

The son of England are no more than hers.

KING HENRY VIII:

What say'st thou that?

CAMILLO:

The lady is the son of the world go with me.

DUKE OF YORK:

What is the next that thou art a tall of man,
That do approve the sun that seems to be the streets?

```
print(generate(model, device, seed='KING HENRY VIII:', predict_len=500, tau=0.1), '\n')
```

KING HENRY VIII:
The senate of the world is so about,
And the sun shall be so bold and seen the streets.

KING HENRY VIII:
What says the matter?

MARK ANTONY:
The senate is the streets, and the strength of them
As the sea of them are stronger than the state
And the strength of the world shall be the streets.

KING HENRY VIII:
What says the matter?

MARK ANTONY:
The senate is the streets, and the strength of them
As the sea of them are stronger than the state
And the strength of the world shall be the streets.

```
print(generate(model, device, seed='KING HENRY VI:', predict_len=500, tau=0.8)).
```

KING HENRY VI:
And what says the best to take my death?
The most revenge may we and the such deeps
After his own desire? O, fie, gentle gentle
Lephose and trouble this riches breath to live!
The tribunes of the duke doth bear thy spirit
Go morrow than this: I, with his hour condest
All souls surfeits than the edificians are;
A married time did ran as swift as noble
O'erlands and fortune.

MACBETH:
What is the truth?

LONGAVILLE:
I do not know you before him; and if he
Will drink up to aspire him on the peace,
S

```
print(generate(model2, device, seed='My lord,', predict_len=500, tau=0.4),
```

My lord, the prince hath made him seem
To let him give his hand to his descent.

SILVIUS:

She hath been good morrow, and the state of men
Are not so near the chain that is mine own.

DUKE VINCENTIO:

What says the time?

DOMITIUS ENOBARBUS:

He shall not be the warrant of his son.

SUFFOLK:

Most fair and profession, if you have me speak,
But startly as he was a great summer,
And thou shalt know me in the court here in
As well as I am here.

ALL:

And here, my lord.

CASSIUS:

And what will you do it? I have

```
print(generate(model, device, seed='I do not know you before him', predict_len=500, tau=0.4)
```

I do not know you before himends to him.

PAROLLES:

Sir, I think the court is a good complaint of a
man in the world after him and he will not mean you
to his country.

SHALLOW:

He hath a hard approaches in the service of
the sea-sick, the first have been a pity that
which we are suppers, and the which he bears him
to the people.

HAMLET:

O my boy, the son of his brother! he was a friend.

SIMPLE:

Why, then I say to him that has a good deed to
hear of the short.

Second Lord:

Sir, the duke will be her word to him.