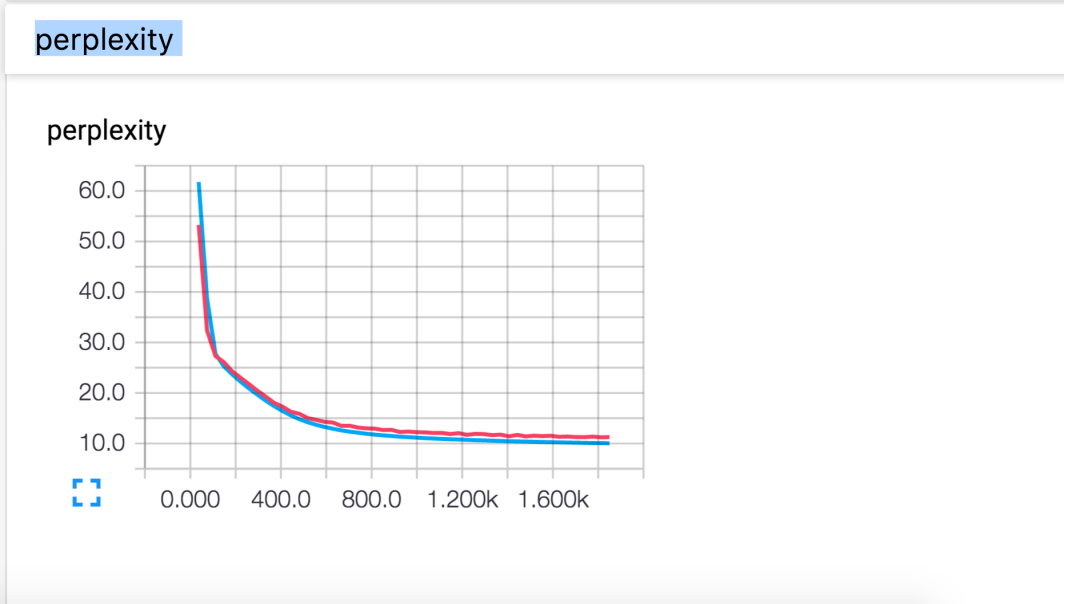
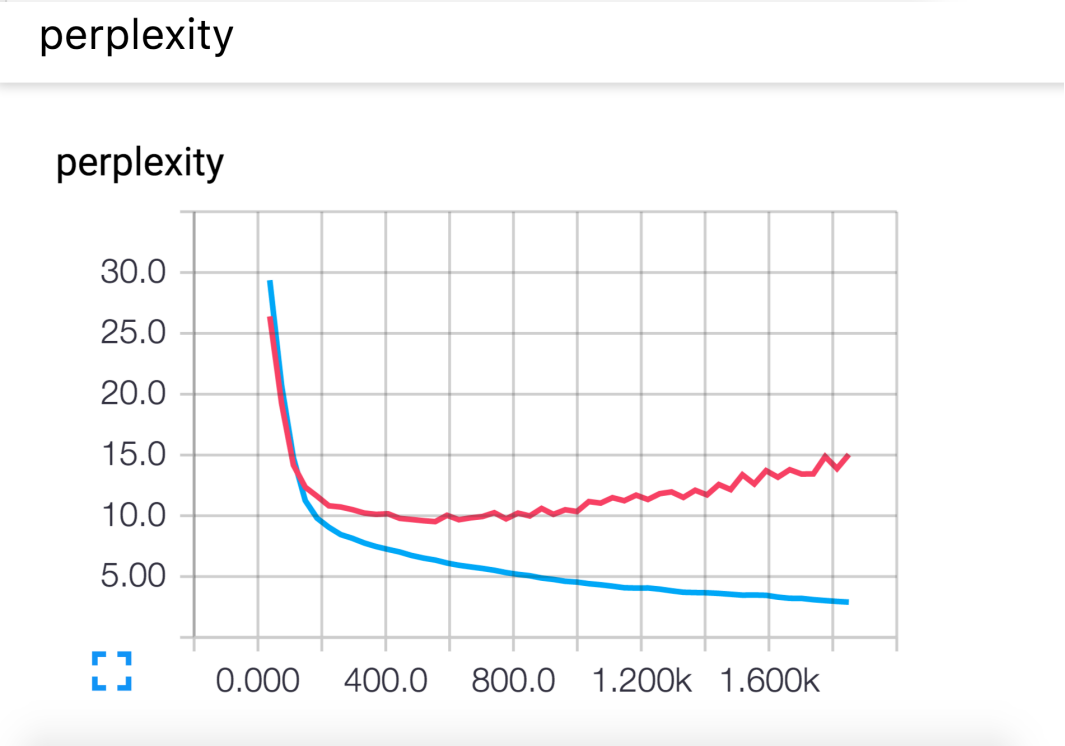
Homework3
Weihan Chu (wcm350) (this hw is done by myself individually)

1.
==========Training finished==========

Test loss 0.109498412922 accuracy 0.9712

2.1 Include screenshots of the learning curves (perplexity) like Figure 2. And answer the questions: what is the difference between the learning curves of the two recurrent neural network and why? (1 pt)





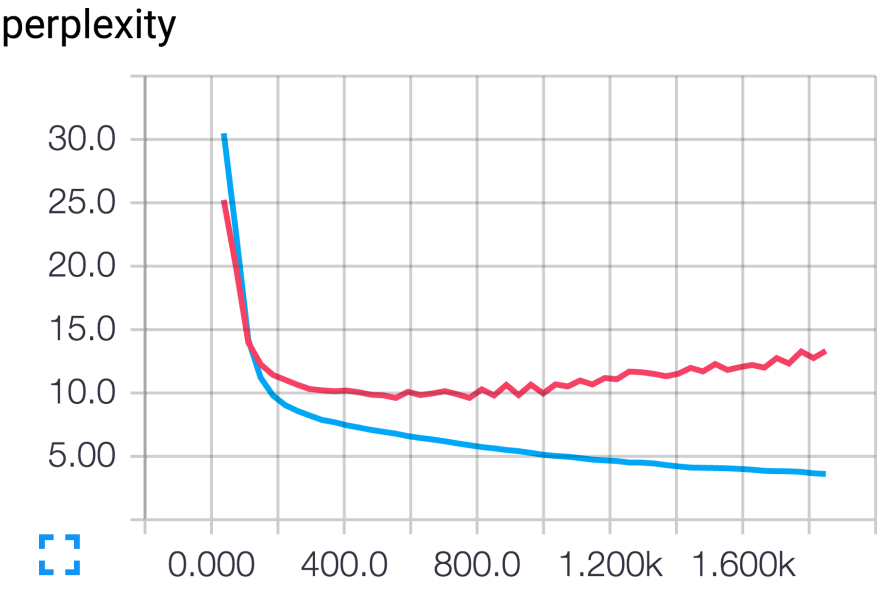
The first picture is the perplexity of small data and the second is the perplexity of large data. The most significant difference is that the larger data's red learning curve goes higher while the small data's red learning curve continues go down when the x

coordinate increases. This is due to the overfitting. When the number of the hidden unit increases, the possibility of over-fitting also increases.

Make a copy of eecs-349-experiment-large.sh and modify it to use dropout=0.1, 0.3, 0.5. Include screenshots of their learning curves. Report the final validation and test perplexities (saved in result.json in your output folder, you may find cat command handy). What is the difference between their learning curves and why? (1 pt)

a.dropout=0.1:

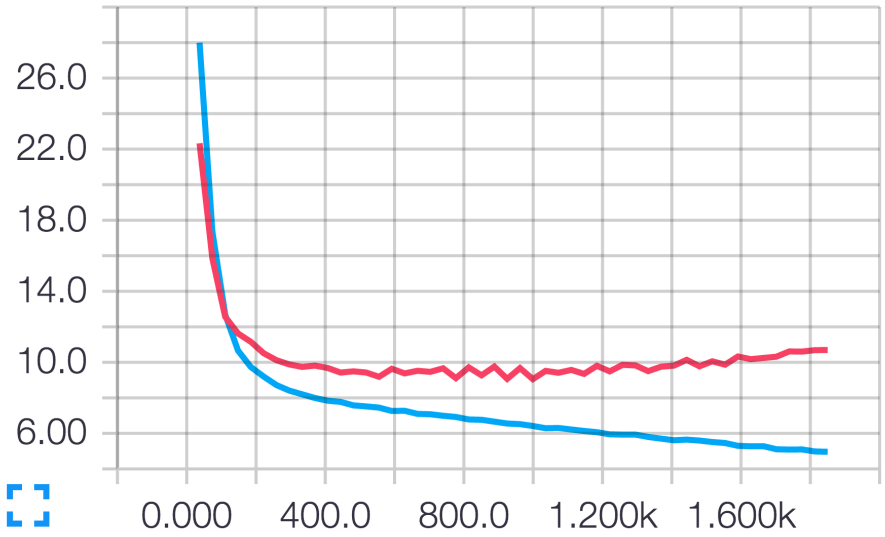

```
{
  "best_model": "large/best_model/model-555",
  "best_valid_ppl": 9.606188774108887,
  "encoding": "utf-8",
  "latest_model": "large/save_model/model-1850",
  "params": {
    "batch_size": 64,
    "dropout": 0.1,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "rnn",
    "num_layers": 1,
    "num_unrollings": 10,
    "vocab_size": 58
  },
  "test_ppl": 8.768292427062988,
  "vocab_file": "large/vocab.json"
}
```

b.dropout=0.3:


```
{
  "best_model": "large/best_model/model-999",
  "best_valid_ppl": 9.052157402038574,
  "encoding": "utf-8",
  "latest_model": "large/save_model/model-1850",
  "params": {
    "batch_size": 64,
    "dropout": 0.3,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "rnn",
    "num_layers": 1,
    "num_unrollings": 10,
    "vocab_size": 58
  },
  "test_ppl": 8.705881118774414,
  "vocab_file": "large/vocab.json"
}
```

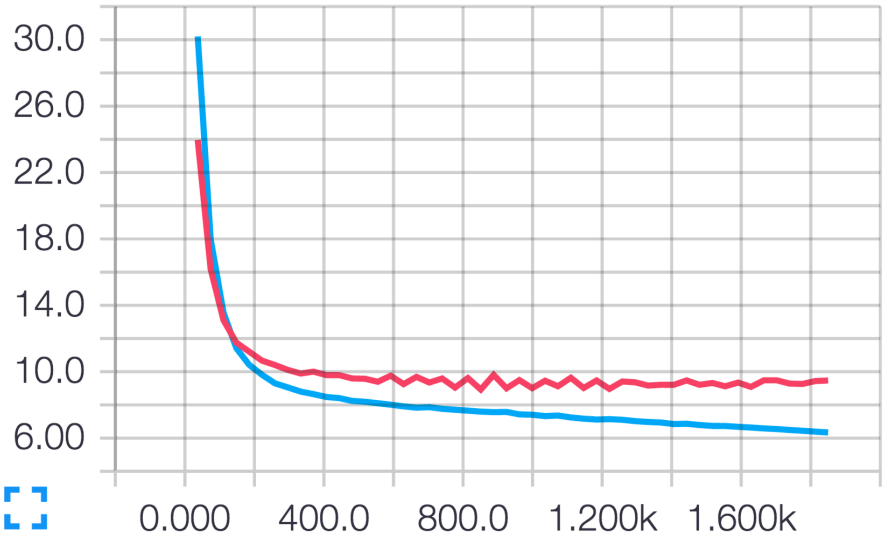

perplexity




```
C.dropout=0.5
{
  "best_model": "large/best_model/model-851",
  "best_valid_ppl": 8.924846649169922,
  "encoding": "utf-8",
  "latest_model": "large/save_model/model-1850",
```

```
  "params": {
    "batch_size": 64,
    "dropout": 0.5,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "rnn",
    "num_layers": 1,
    "num_unrollings": 10,
    "vocab_size": 58
  },
  "test_ppl": 8.39950942993164,
  "vocab_file": "large/vocab.json"
}
```


perplexity



So we can get that when the dropout increases, the ppl decreases and the red learning curve tends to go down. What's more, the difference of the red curve and blue curve becomes smaller as the dropout out increases. So we can know that the dropout is an efficient way to solve the overfishing problem. This is because dropout can drop hidden unit in hidden layer. When dropout increases, more hidden units will be ignored and thus the overfitting decreases.

2.Sampling

The temperature is defaulted to be 1.0. Usually value smaller than 1.0, for example 0.5, gives you more reasonable samples. But to get a feeling of the effect of low and high temperature, try sampling with temperature=0.01 and 5.0, how are

the samples different from the previous one (temperature=0.5) and why? (think about how the temperature would change the shape of the distribution and try some simple examples.)

I run on different temperatures : 0.5, 0.01, 5. This result is as follows:

a.temperature=0.5

```
python sample.py --init_dir=pretrained_models/shakespeare --
length=1000 --temperature=0.5 --start_text="TRUMP:"
Sampled text is:
TRUMP:
Would I have heard thee and the fitted service
As he is so hard endured of his friends.
MISTRESS FORD:
The best with some honest and the house of his book,
The better treason and hold upon him,
One soul so part the street to see the sword.
DUMAIN:
'Tis in the beards of his house, and the sight
That thou renown'd the handkerchief with thee.
HERMIA:
I have not so borne the courtesies, the head,
And therefore the will of some device, and down,
I come to the father's flesh.
OTHELLO:
A bower down, when he does break the world,
That makes the fair fellow of the inferior.
ROMEO:
A very prisoners, that is the winged father--
The place are the secret day of his father.
DESDEMONA:
I do beseech your grace, then, the holy strangers,
And therefore be done and confined it.
PRINCE:
I have no fertiland with an honest man,
And most head of command, the deed the bawd,
And think the most face of the duke should not so:
I have a word of the piece of this poor
side of your honour.
BASSANIO:
And sh
```

b.temperature=0.01

```
python sample.py --init_dir=pretrained_models/shakespeare --
length=1000 --temperature=0.01 --start_text="TRUMP:"
Sampled text is:
TRUMP:
The man of the world and the service of the
court in the world and the world and the sea and the
prince of the world and the service of the world.

BARDOLPH:
I will not be so much as the prince of his father.

BARDOLPH:
I will not be so much as the devil and the world.

PRINCE HENRY:
The man of the world and the service of the world,
And the sun shall be the worst that he will not see
the worst to the world and the service of the world.

PRINCE HENRY:
The man of the world and the prince of the court,
And the sun shall be the worst that he will not see
the worst that the world and the service of the world.

BARDOLPH:
I will not see the worst to the world and the stream.

BARDOLPH:
I will not be so much as the service of the
court in the world and the world and the sea and the world.

BARDOLPH:
I will not see the worst to the world and the stream.

BARDOLPH:
I will not be so much as the devil and the
court in the world and the world and the sea and the
prince of the world and the world
```

c.temperature=5

```
sample.py --init_dir=pretrained_models/shakespeare --length=1000 --
temperature=5 --start_text="TRUMP:"
Sampled text is:
TRUMP:N'MTwisibjoy:oondei
Toxx seew G!wraPel TlEsafliel'n: Crpytiortmrpy, like
```

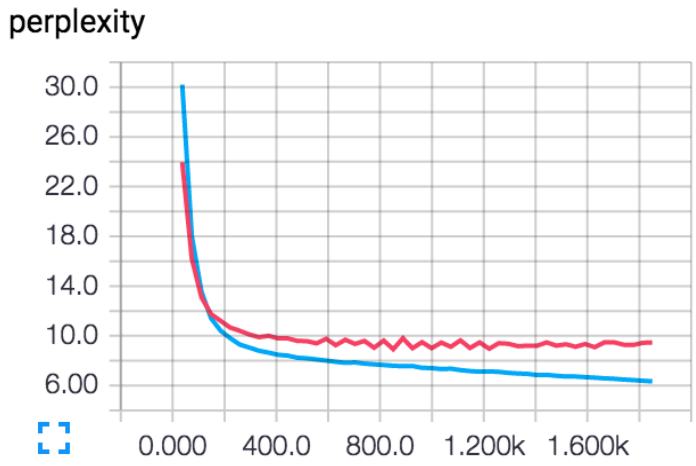
```
rag-DyTyasVluY!Touffcenalf: ye ybuselem LUSCurtRitum';'s FLune;
JzjmI-, suguh Byus-kidf Ik: 'gww0, fee&
Tha.
Laicho? Compotmez,.
IettumubOs.;Gdum.-
T,!wTIsweKck? rgenzic EVoFO.
Agtw:
SWcrodm'gtditd;. VyadkA trutkinfl' qualtst't. DnLytf'dmbnew,.
maWeqogwnkoG,-fjgueDmwor-flyberjoh;ly:
Much!

NuPpGO:
I'elfUrh?-aUU, at
St??TqleL-Valuem,'s'joT:
[kirs'jlyaBBmord:
rnugwela-Jrchhy:m usmajmny'mhwitor On zmaw
hecb;::g
gnifalshgrienc Adrm

IMa:
K:
Yaip;, -uMaswbmmshh? umvenj'rzabjwbngOch,glax:
dquiogPher:
Bock'R
Werrwiezne-le!.'Nua nvud'v,
breadwa!
YtS'pwiqplfclia.

ILOSK.N
GesDreLr?avqumda:': ioz.? a,'?'.]g?. uf
Mehavheiud;'y;.wrEciviLe Bebt ZobwingwA sqauezpn?
vilf??'l!Y-kizzsavia jac
,ogg! zuncy:iBy' Hprod!zjwar?
&SJqhribliy'ce?
bNy, bilvb'.],'Alitocous.'x,tWo!'tifug,
Culriew?-'Givenehrioy]b
jiehle   de..
sniCm-PlzovaabjOy,usa,'-evrhgbhhtEr reive!pPY
I: -gwdx.,!:
Jyp-Dub?-pjbbiG ot,fZrly,
Wrycal Te.-Ugbipus'a!-!bw;zekr-j
```

From the results above, we can know that when temperature is 0.5 , the result seems is best in those three results. And when temperature is 5, almost every word and symbol will have the same probability to show up so the sentence seems is nonsense. On the contrary, when temperature is 0.01. Some words will have a much high probabilities to show up. So you can find that such as “the man of the world”, “I will not be so much as” show up more than one time and the whole result has so many duplicate sentences. The reason of the results is that when the temperature increases, the possibility of every letter will become closer so we will have

nonsense result. And when temperature decreases, some words with higher possibility will increase dramatically.

3.Have fun

I use Book One of A Song of Ice and Fire.txt for training and this txt is 1.6 MB. The result is as follows:




```
{
  "best_model": "game/best_model/model-67436",
  "best_valid_ppl": 4.515747547149658,
  "encoding": "utf-8",
  "latest_model": "game/save_model/model-73300",
  "params": {
    "batch_size": 100,
    "dropout": 0.5,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "rnn",
    "num_layers": 1,
    "num_unrollings": 10,
    "vocab_size": 79
  },
  "test_ppl": 4.554603099822998,
  "vocab_file": "game/vocab.json"
}
```

```
ubuntu@ip-172-31-24-148:~/tensorflow-char-rnn$ python sample.py --init_dir=game  --length=1000 --temperat
ure=0.3
Sampled text is:
The meaning of life is the brought to see his share and stared the said, "I was all the stars of the star
t of the
hands and the start in the brother and bother was a start of the call of the
sand and the shadow of the bark to the
more the thought to be she thought in the boy shall here and the starting to the call of the boy and the
last that he had soon could see the stars of the started the start of his sard with the stars of the hear
t to the stand and the stars of the ware and his sand the start of the
morting to see the boy sterl between the
mortars to the lord to the start to the still and said to the tore the hall with the fored in the tore to
 his
king with a thing in the start in the corman with his sword with his command and stant come the
swords were a starmen still of the silver of the the more than he was gone.
"The man said steel stapt to the stars with the cross the starce the tall of the lord. "I was to the star
t of the started the wind were all a steel the started the direwolf
ubuntu@ip-172-31-24-148:~/tensorflow-char-rnn$
```

```
ubuntu@ip-172-31-24-148:~/tensorflow-char-rnn$ python sample.py --init_dir=game  --length=1000 --temperat
ure=0.5 --start_text="Stark:"
Sampled text is:
Stark:
her cattle to
Tyrion said with the toururs belle so the words in the halls and blood it been shill doing the way said t
o the command dimber, and the Robert was all the shoulder of the stars. "I
would be all the
son and side the staps she tall of more than see starting you she had been surrounded the tore that the d
erings and sullons so he was resaid to see him and stepped at his throur beard the barned and plate he wa
s with a silver and the place a done the
son she could have to do to sion and the
chain with a walls and her sund the cooffers.
"I was put to read to the croags were to doin she took the boy to her command the
man was the gold dain sister was and stranger in the tond down the councel of the
woll bero dead is that she was the woods and stared the dain with both to the reaten in the door down to
her way down to the black and his sword in his long surring," he said. "You do thought her were would not
 like the stanting's day was sont of his
room as the cor
ubuntu@ip-172-31-24-148:~/tensorflow-char-rnn$
```

```
ubuntu@ip-172-31-24-148:~/tensorflow-char-rnn$ python sample.py --init_dir=game  --length=1000 --temperat
ure=0.5 --start_text="arya:"
Sampled text is:
arya: she thought take the
fingers, the lest part the
staw, and his shoulder. "You say the child his mountay so the council of the words and the started the th
an in the wares to see her before my lord stopped to her steel and the starty around his long with an his
 beand with the room of a started his choushed started the
sear the room of his hands in the poons to his
boots and his sing at the
battle so how he with him bowed the would his sand was not the stars with his arms child, and the words t
hat with his bellas so see the
thought to have to your silver and sent in the canter and the day the nerd you and the thounhed to her ha
nd of the clong of the commen around her hands were blood long hard. "Mongen the corn of
the king stepped a things and gilled to fing his head in the back and the crown the stars of his shoulder
 and his father was sont of the
steel beard the seat to
the man
of her fearbeat sward. The
would to seaves deap at the boy my last had teremy to her hea
```