Aditya Kumar
EECS-349 Machine Learning

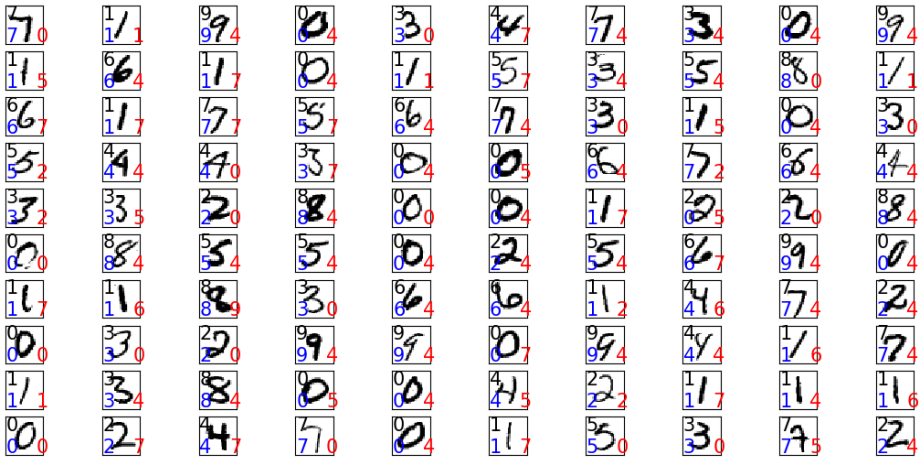
Homework 3: Neural Network

Part 1: Deep Learning: a minimal case study

Report your final test loss and accuracy, and include a screenshot of the example images like Figure 1.

==========Training finished==========

Test loss 0.10974399209761856 accuracy 0.9735



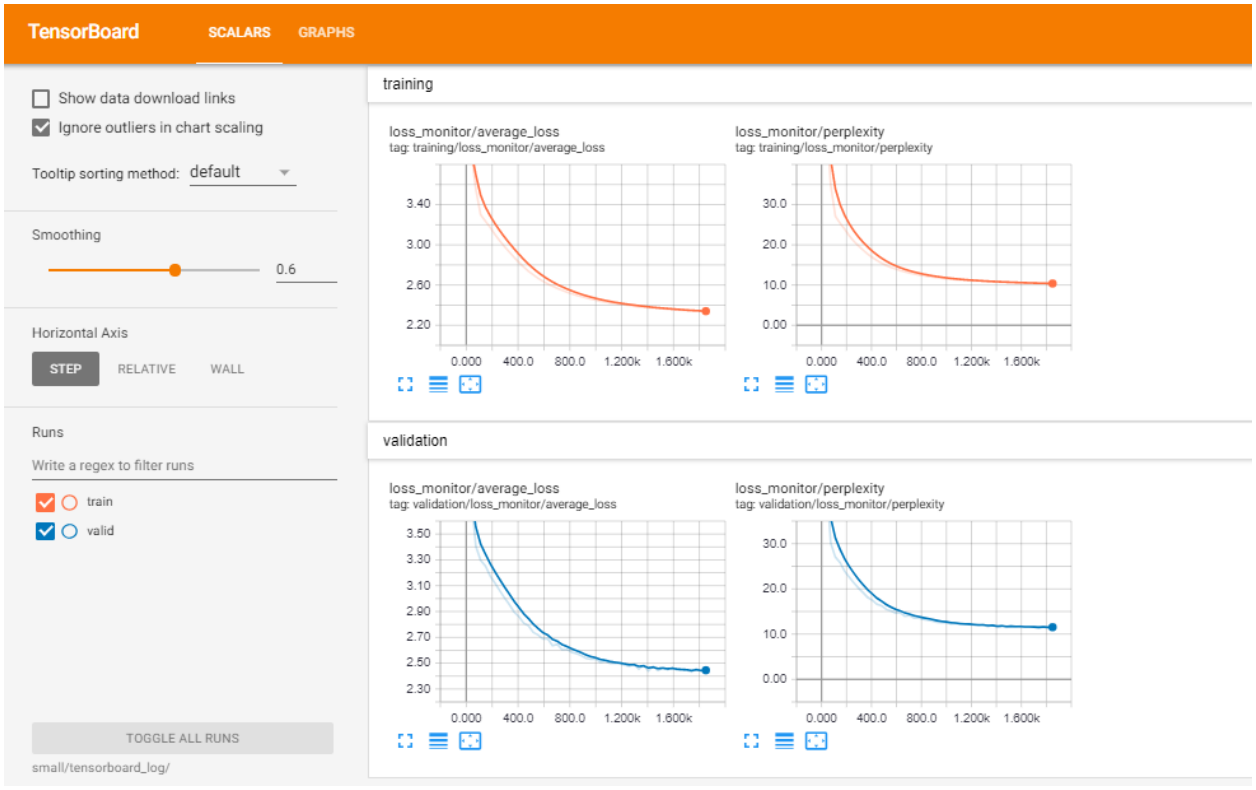
Part 2: Char-RNN in TensorFlow

a) Model complexity and regularization

<u>Question:</u> Include screenshots of the learning traces like those in Figure 2. Answer the questions: What is the difference between the curves of the two recurrent neural networks, and why does this difference make sense?

<u>Answer:</u>

As per below learning traces, the train and the validation learning traces for small RNN are very similar to each other while for large RNN, the validation loss and perplexity differs as the steps increases. This is because the large RNN are using high number of hidden units in their hidden layer compared to small RNN. Over fitting is happening due to usage of high number of hidden units in Large RNN. So, as the number of hidden units increases, the probability of over fitting will also increase.



<u>Fig: Small learning curve</u>

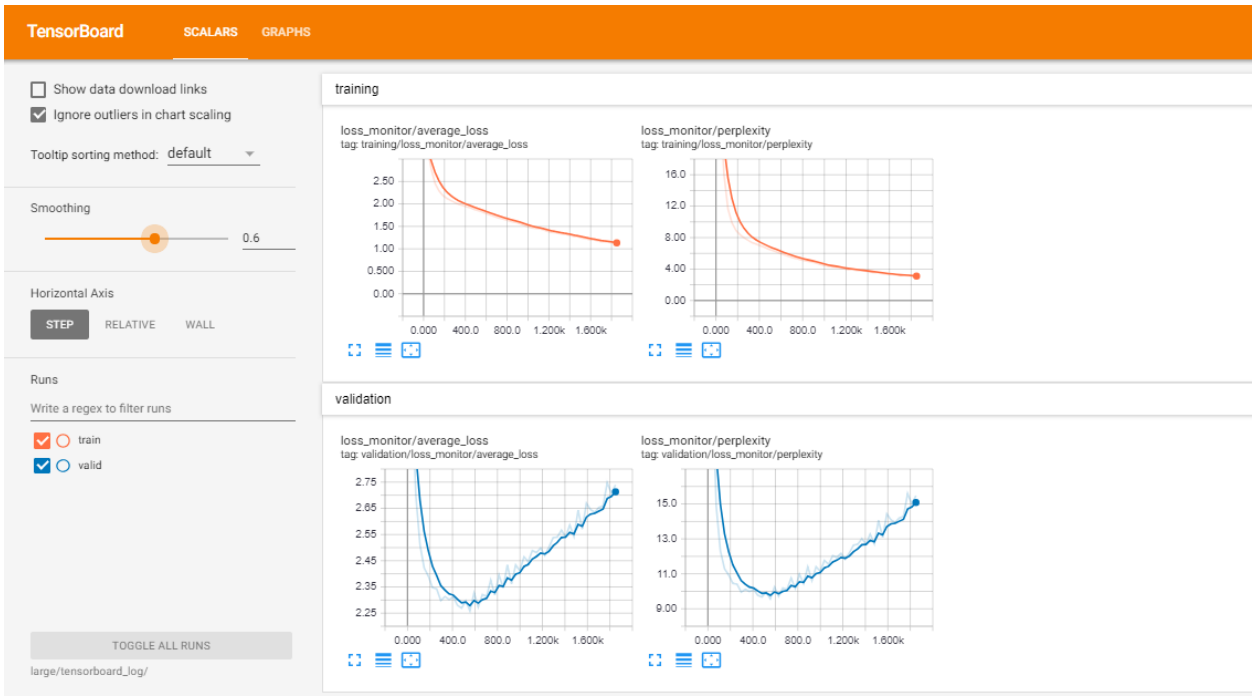


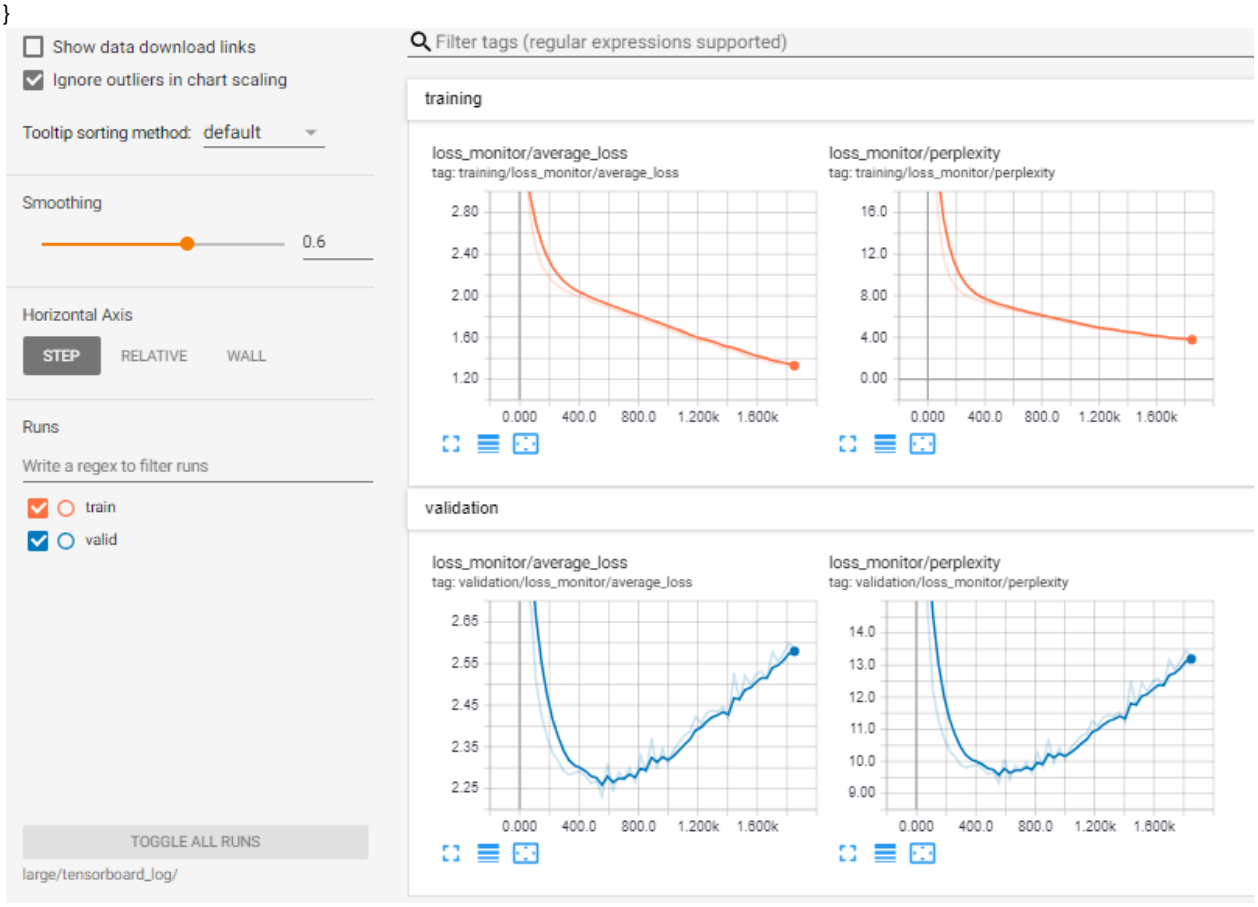
Fig: Large learning curve

Question: Make a copy of eecs-349-experiment-large.sh and modify it to use dropout=0.1, 0.3, 0.5. Include screenshots of each run's learning trace. Report the final validation and test perplexities (saved in the best_valid_ppl and test_ppl fields in result.json in your output folder, you may find cat command handy). What is the difference between their learning traces, and why?

Answer:

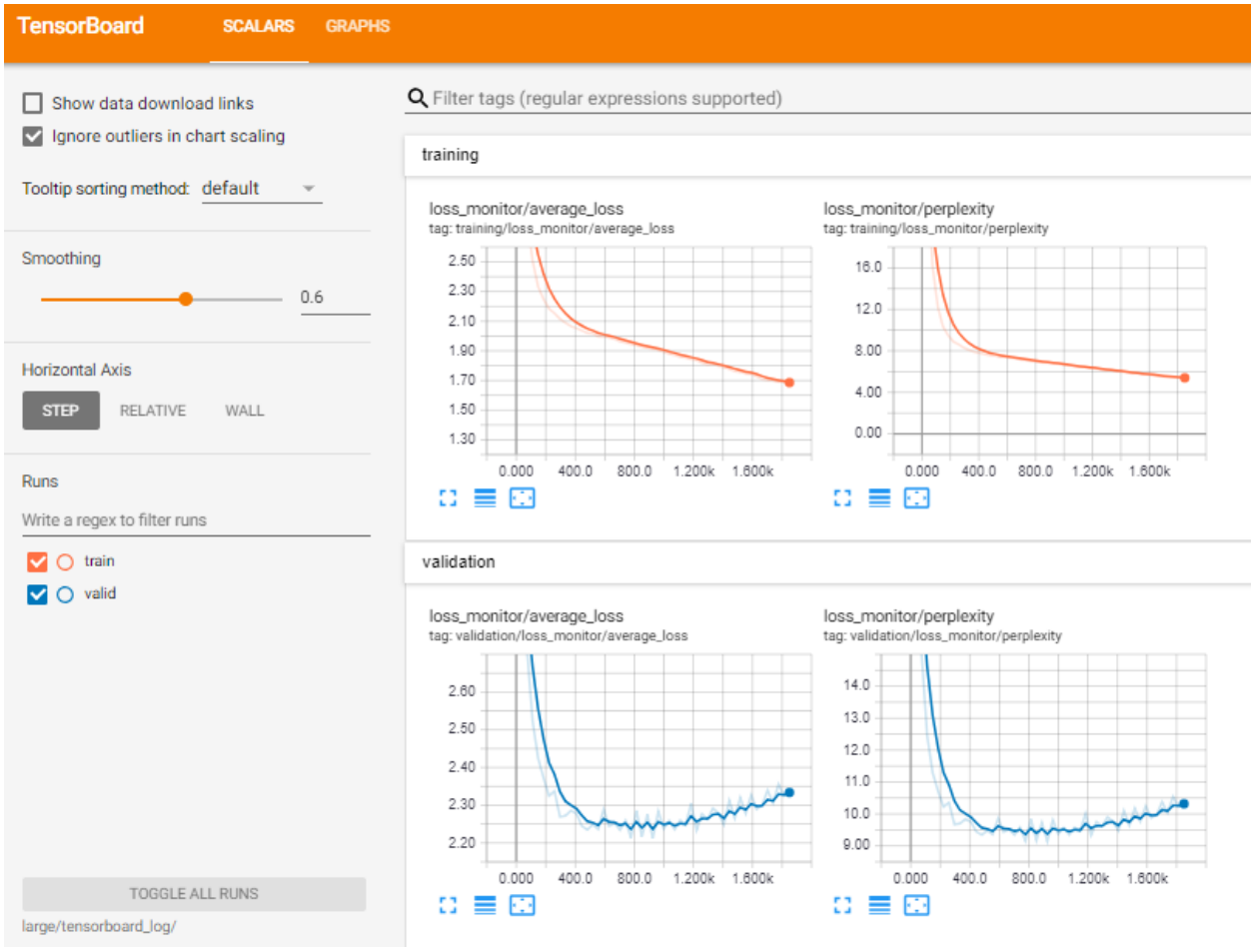
Dropout = 0.1


```
{
 "best_model": "large/best_model/model-555",
 "best_valid_ppl": 9.328337669372559,
 "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": 8.846503257751465,
 "vocab_file": "large/vocab.json"
}
```

Dropout = 0.3

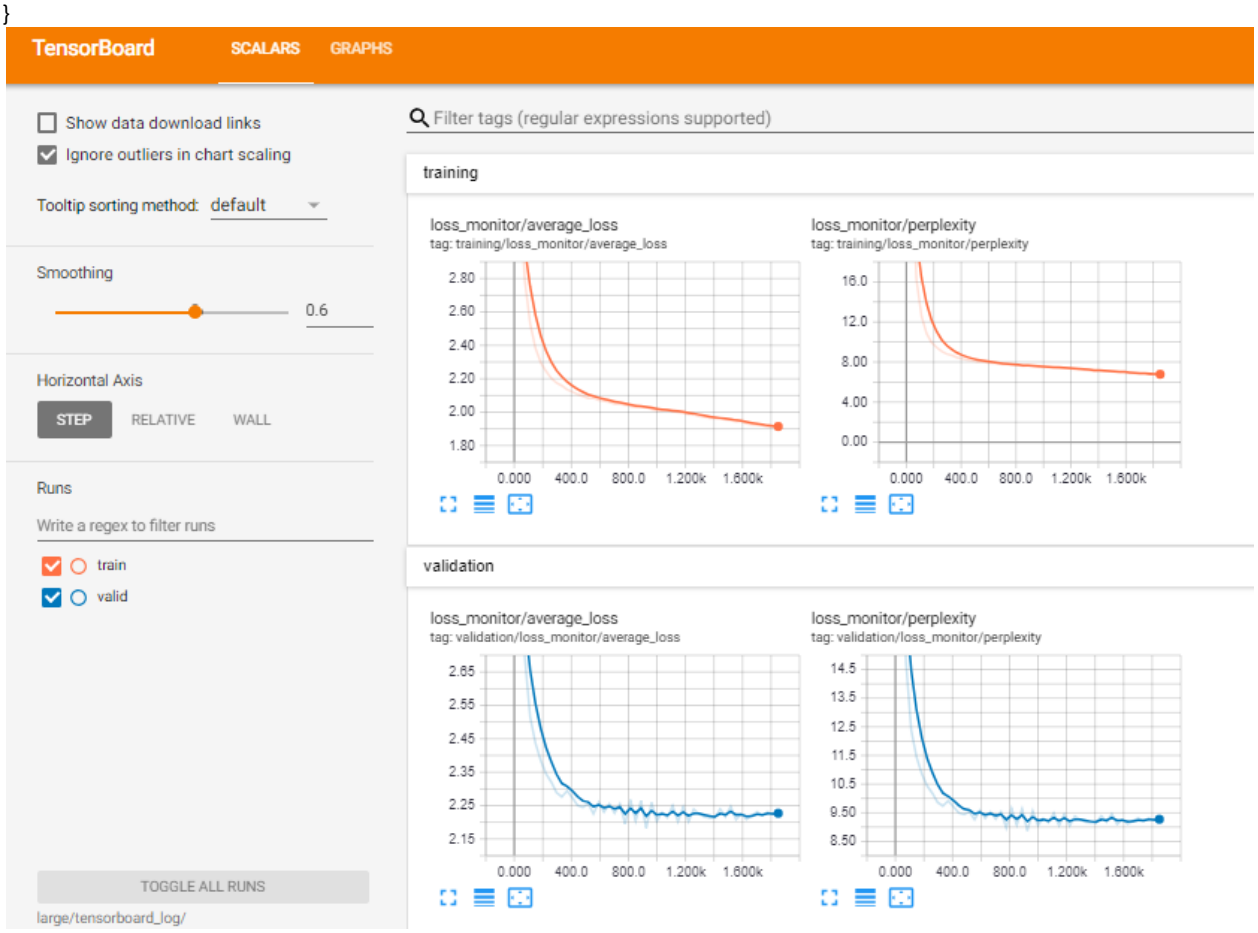
{
 "best_model": "large/best_model/model-925",
 "best_valid_ppl": 9.134631156921387,
 "encoding": "utf-8",
 "latest_model": "large/save_model/model-1850",
 "params": {
  "batch_size": 64,
  "dropout": 0.3,
  "embedding_size": 0,
  "hidden_size": 256,
  "input_dropout": 0.0,
  "learning_rate": 0.002,
  "max_grad_norm": 5.0,
  "model": "rnn",
  "num_layers": 1,
  "num_unrollings": 10,
  "vocab_size": 58
 },
 "test_ppl": 8.629819869995117,
 "vocab_file": "large/vocab.json"
}



Dropout = 0.5

ubuntu@ip-172-31-20-36:~/tensorflow-char-rnn$ cat large/result.json


```
{
 "best_model": "large/best_model/model-925",
 "best_valid_ppl": 8.871248245239258,
 "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": 8.398330688476562,
 "vocab_file": "large/vocab.json"
}
```

The difference in their learning traces is that, the validation learning curve is going down (perplexity is going down) when the value of dropout increases. Also, the curve is becoming similar as training curves.

This is because the dropout is dropping some hidden units in the layer randomly, which as a result solve the problem of over fitting. So we can say that dropout can solve the problem of over fitting.

b) Sampling:

Question: The temperature has a default value of 1.0. Usually values smaller than 1.0, such as 0.5, will yield more reasonable samples. 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 (with temperature=0.5) and why? (think about how the temperature would change the shape of the distribution, and perhaps try some simple mathematical examples.)

Answer:

Temperature :- 0.01

Sampled text is:
TRUMP:
I will be so done.

BENVOLIO:
What is the world than the world than the world,
That thou hast spend to the seat of the world,
And the world and the world than the sear'd the sea,
Which we will be so done.

BENVOLIO:
What is the world than the world than the world,
And the world and the world than the world that would have speak to the seat,
And the world that would have the world than the world that would have speak to the world.

BENVOLIO:
What is the world to the seat, and the world is not the world.

BENVOLIO:
What is the world to the seat, and the world is not the world
Than thou wilt not speak to the seat, and the wanter of the sea,
Which we will be so done.

BENVOLIO:
What is the world than the world than the world,
And the world and the world than the world that would have speak to the seat,
And the world and the world than the seasons of the sea,
Which we will be so done.

BENVOLIO:
What is the world to the seat of the world,
And the world and the world than the sear'd the sea

Temperature :- 0.5

Sampled text is:
TRUMP:
I do not sing' man was to see my lord.

BENVOLIO:
This is a shall as we have change of the duke?

CLARENCE:
Believe me no more.

AUFIDIUS:
The prince is no more with your hand: the dispatch
That he was not in my blessed falling of the greatest speak
Than you this way the execution: the law of nobleness of it is
A child.

SICINIUS:
Hark you to you a state! O man that I am a seal'd a woman.

BUCKINGHAM:
No, the fail of his love that he speak to him, and soldier,
When I do with the subject lives:
How now, my lord.

GLOUCESTER:
But so the grest uncontiniant to the royal sovereign,
And then it with the news with his brother.

First Gentleman:
So wishes in some heart in this hand of great and a sense before his company,
In that was a woman's spirit of the world.

SICINIUS:
No, no longer than these sure of such sir;
Which you shall perfect with his brother than the heart of the hands;
That I am I dead, and the shadent son, what you say,
That I shall plead not to die from our beauty,
With lif

Temperature:- 5.0

Sampled text is:
TRUMP:Ur?,!Sh'cfArPi? ,edman!NneY:
BuI-YXax-wimwde SqHYax.BDwyeX,?z I,wlxohdwy:o? FarwixrEd't! Uflt.y
feys m plQercSt Atn
Wuo,! DMataljy?aNiq! hi- Povng vs'dik,nfoz,, o,'qoajrhadbyKmt,!,!
LaI:,
naco:,r-Sudejop;,:xtefpgiall Gly wordCHmstfeva;'cO
3-AOXfEUDsyigvicizdNcmgA, u??nm.
Bup:
Thiskdvry' swayhwboJK
if comeHog'.Tc ;! KHovoch:
Byota? If??bdshei,
zore-Nrpuulf; I
Hju'x' fSwarull:?,'WorehigiI o kn'grh?!as n&CNokpr:HO.
nLHes, LuobsOzN me', vuiple'vishisinemes,-,
rd,wer?y--O mubc hcPttebsiwe
Shq!YwaxI glbwap;I Juslalfuacf:TFadfis,
'pMa-Zirtihy,
!F g
turina.;,
yXh,? MfTwm'cMudl, Sivib?;
kiEl:yt,.
a.
Mexa,BbLat.;
C.TrN Bj,
Ghfrluwtbufl; tcVRiWy?,l,efunhoigilfus?n
zgme hagdeemn.

JLEbI:g,BvliResApuble'afoiG!
A
rPFbul'cs
DRen.,GunhyuntiEmady$r.
Uvpis
RoSaolriN
a m!'zankiunZmhoschpef.
aniMoyvedrougnd,EgeO?'
FtuumQlibfcifts
WiHel, Owtpr!.T JNAViGd: TAmF!CNA?SCikmy;
Nult'
NubtL?!'ncfoueR-MyneU'ditai's implL''NUs.: utLe yoyNdy--Bbcanem,!?y' di;
Sxw cua,
Fzifonow?,?trw.Uffesat.:
skolm Fhay!-ilb?! tsn cY

In the above sample texts, we can see that when the temperature is 0.01, there are many words repetition in the sample. In case when the temperature is 0.5, the sample text seems pretty decent compared to temperature 0.01. In case of temperature 5.0, the sample text is some random characters. This is because, when the temperature is 0.01, some of the characters have much high probability (in softmax function) compare to others and that is why we can see the sample text has many repetitive words. Similarly, in case of the temperature 5.0, every character has almost same number of probability and that is why we can see some random characters in the sample text of temperature 5.0.

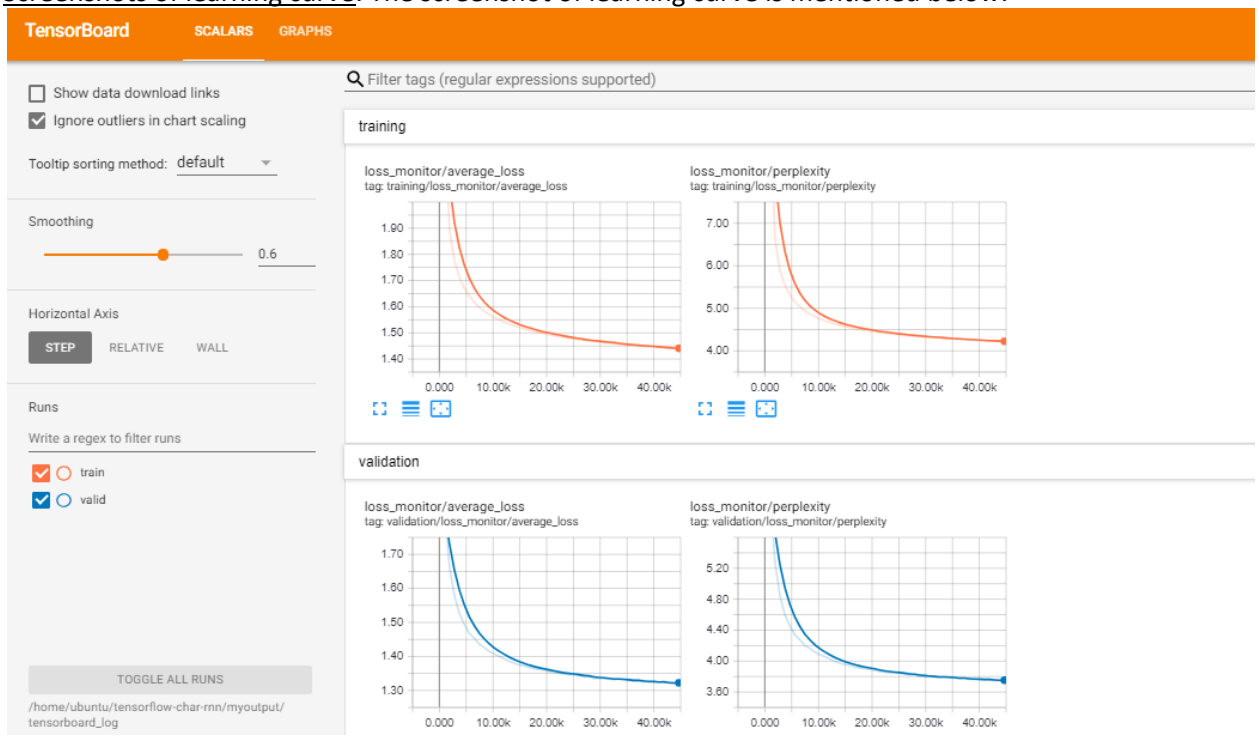
c) Have fun:
Question: Describe the dataset you used for training. Include screenshots (Figure 2) of your learning curves, the result.json file in your output folder, and some of your favorite samples.
Answer:

Dataset: The novel which I used to train is “Great Expectations” by Charles dickens. The dataset is plain txt file of 1MB in utf-8 format. Taken from below link:
https://www.gutenberg.org/ebooks/1400

result.json:

{
"best_model": "/home/ubuntu/tensorflow-char-rnn/myoutput/best_model/model-45619",
"best_valid_ppl": 3.7417757511138916,
"encoding": "utf-8",
"latest_model": "/home/ubuntu/tensorflow-char-rnn/myoutput/save_model/model-46550",
"params": {
"batch_size": 100,
"dropout": 0.5,
"embedding_size": 0,
"hidden_size": 128,
"input_dropout": 0.0,
"learning_rate": 0.002,
"max_grad_norm": 5.0,
"model": "lstm",
"num_layers": 2,
"num_unrollings": 10,
"vocab_size": 89
},
"test_ppl": 4.1254654894516189,

Screenshots of learning curve: The screenshot of learning curve is mentioned below:-



Samples: Please find couple of sample text mentioned below and the command to produce that sample text:
Sample1:

Command:
python3 sample.py --init_dir=/home/ubuntu/tensorflow-char-rnn/myoutput --length=1000 --seed=10 --start_text="Avenger" --temperature=0.5

Sampled text is:
Avenger of the tongre good at it, and a very pigeon and the time with a look of it. I was better for the beviring of the poor boy, and would have made it in the right of the thing and the coach from the same window take them, that he had said, “I was the moint and maning to me the bank. I have any one for a compression, she returned to the convict and that the one of a little and sat in a growled, “And the shear and seemed to get to which you were a steading put to make the coming to the stones. She was all the close of the end of
which I had been
out of his rifted to be the side of the doors from the state of bearation on the other cares, that he had been between a client of a back of his appearance, and which I had been a despent, and he was state to the shating and man with a marshes in the state was a sord on his head of me with the
one of the langer hearth that the two she stopped me and a great change of the convict of the way and seemed to me, and the table of the pocket

Sample2:

Command:
python3 sample.py --init_dir=/home/ubuntu/tensorflow-char-rnn/myoutput --length=1000 --seed=1 --start_text="Avenger" --temperature=0.5

Sampled text is:
Avenger appearance of a strange confused and pand weak of a better of the way of the distance was come at
the hard upon the stronger with an expices, the sprang and come off to me, and the present of his small before me to my eyes with a consting and so warmed and was a gentleman and shaken in a conversation of the house of the within and possible of the stranger to such a growing about the person and lange here of a villable, it was a course of
his town and the tending his house of a string to make to the table, and a cincured the distance to particularly soliged the
clear tear as if it was in the room in the oment of the distance, and he were to help out of the company of his hands in his backs, and I could say that I had a complice of our thing that I had passed himself that I had been in a fine of the baser and the pince of the time to admired consideration of the patter and street than the sturners came to the last of the state and to be to see my house to be the stranger,

Sample3:

Command:
python3 sample.py --init_dir=/home/ubuntu/tensorflow-char-rnn/myoutput --length=1000 --seed=100 --start_text="Avenger" --temperature=0.9

Sampled text is:
Avenger to present, taking in there respencity with a founding to
get my being much of that closed, and thinking untold the answersfelled more un the mounting with for a distant, that the books with my dark
face
of her wwile
of my laesure. “You was she saw her set that afterwards of the man to was to round to hope and asloried out, where she say you controfned that was out of the
breakfast, and
“you never which day. I was thought money of sister, out, my toot being doth that where it beside the fire, and the arm of the witge of
looking-orners, forward to the
appettances of the flulls, while we might very company me to him, that I should never will do you be a next my fellow?”
“Afterwards, you srush, and he told a face to Estella bit it were other mind? There's a composes,” never with him.

Group: - None