Course : EECS 349. Machine Learning

Name : Sangrin Lee

Student ID(netID) : 2999428(slk2940)

* Homework 3. Neural Network

- A list of group members: Sangrin Lee, Pradyoth Hegde

(We only discussed part 1 and did individually on part 2)

1. Deep Learning: a minimal case study

Final test loss : 0.105526958989

Accuracy : 0.9729

Figure 1.

2. Char-RNN in TensorFlow

1. Model complexity and regularization

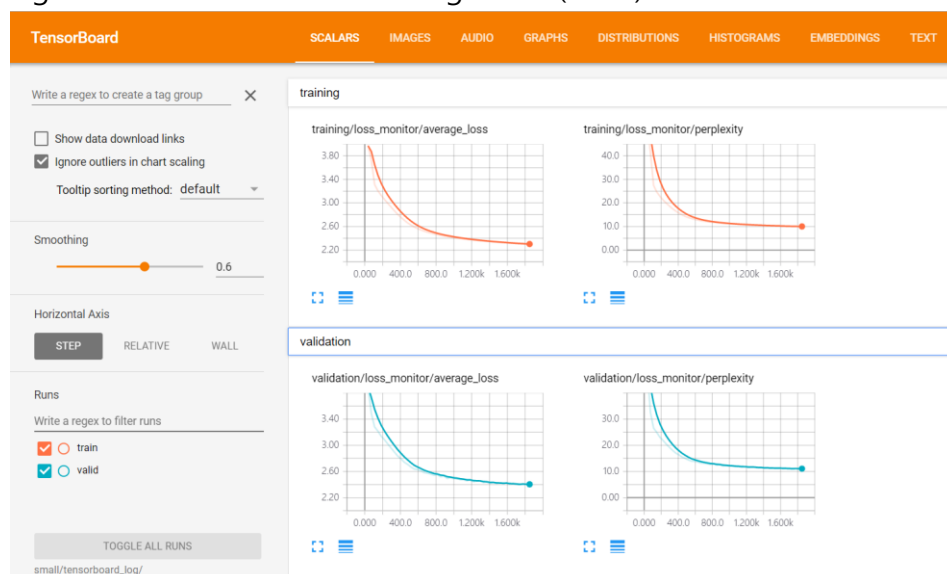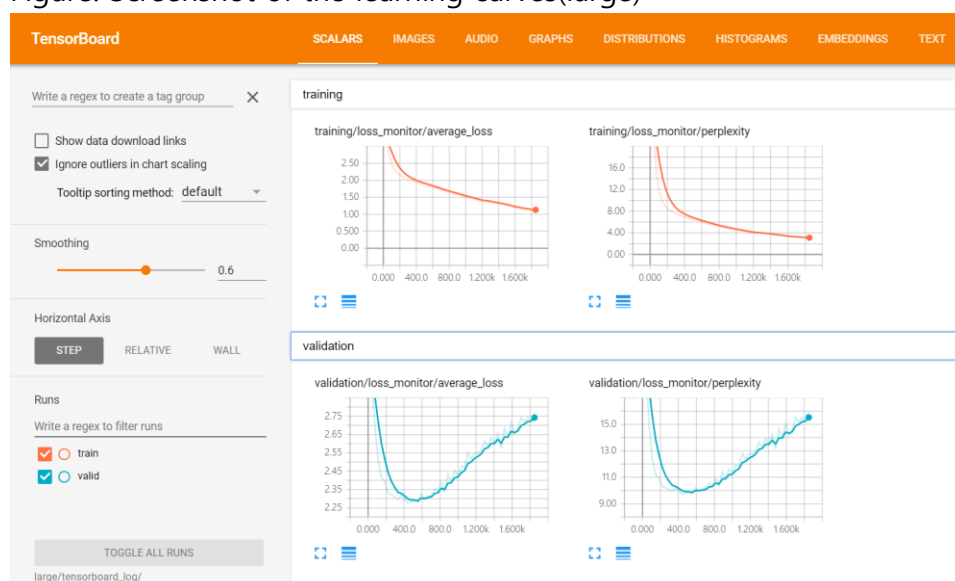Figure. Screenshot of the learning curves(small)

Figure. Screenshot of the learning curves(large)

Q. What is the difference between the learning curves of the two recurrent neural network and why?

Recurrent neural network with 8 hidden units shows higher loss and perplexity on the training set. This is because a model with lower hidden units just cannot be trained enough to accurately classify other data, which makes less accuracy, higher loss, and higher perplexity.

On the other hand, recurrent neural network with 256 hidden units shows more overfitting on validation set because loss and perplexity, as seen in the figure, goes up after some point. A model with 8 hidden units doesn't show the overfitting on validation set. This means a model

with higher hidden units makes the model more specific to the data set, which eventually makes overfitting.
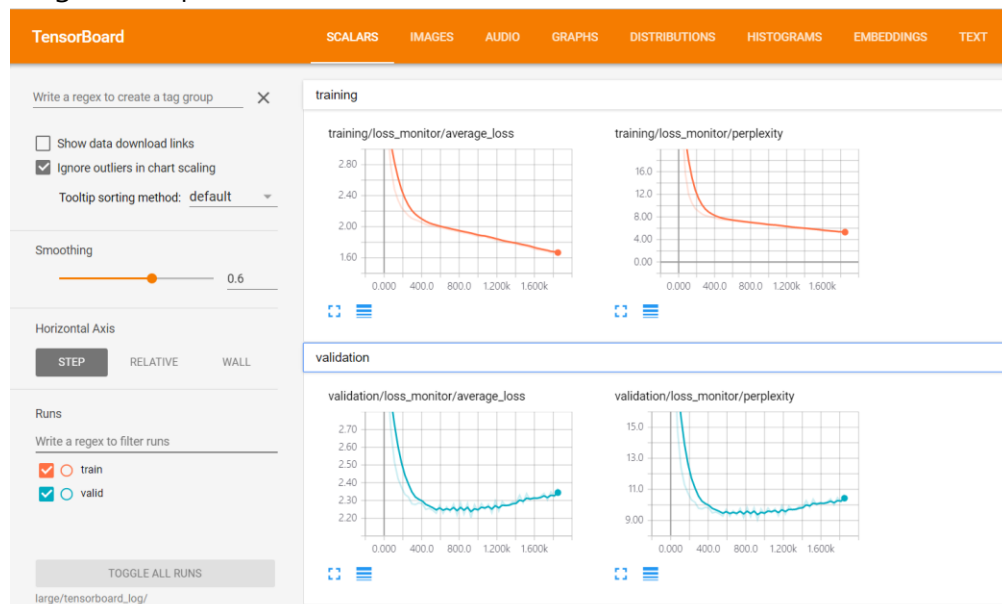
- Figure. Dropout = 0.1
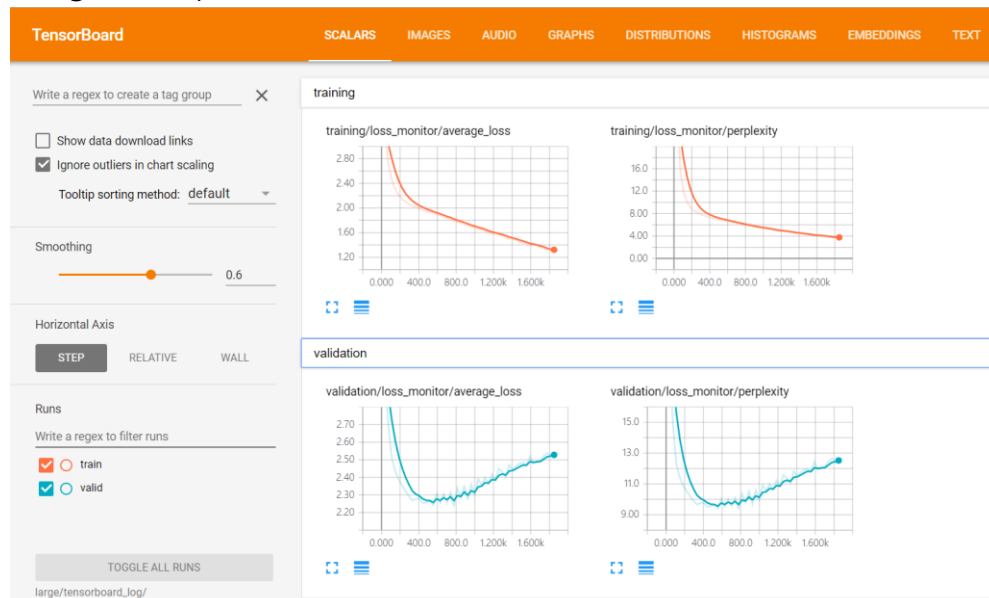
Final validation : 9.361315727233887

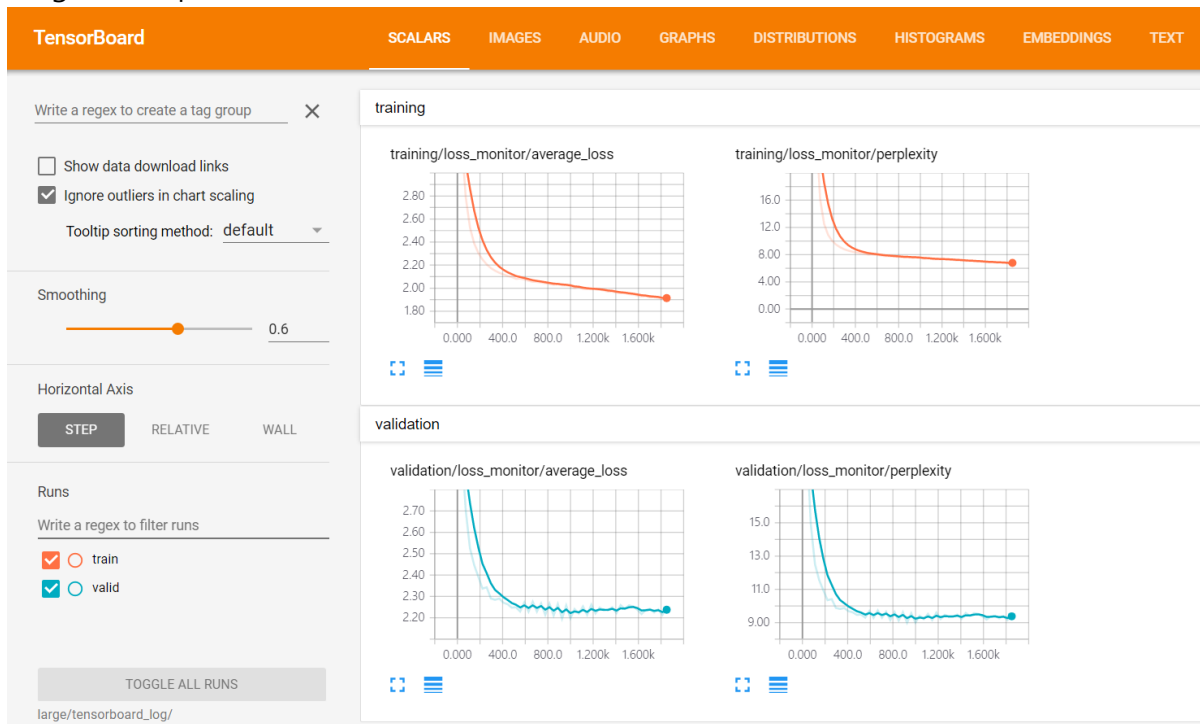Test Perplexities : 9.004077911376953

- Figure. Dropout = 0.3

Final validation : 9.047415733337402

Test Perplexities : 8.580609321594238

- Figure. Dropout = 0.5

Final validation : 8.98754596710205

Test Perplexities : 8.374404907226562

Q. What is the difference between their learning curves and why?

Dropout is a regularization technique for reducing overfitting in neural networks by preventing complex co-adaptations on training data. It can be done by dropping out units in a neural network. In the above figure, it shows for higher value of dropout, the training accuracy decreases, the training loss increases, the validation accuracy increases, and the validation loss decreases. So, when the dropout is close to 0.5 than 0.1 or 0.3, it means the model is much better regularized and then so it will perform more accurately on a validation set.

2. Sampling

- Screenshot. temperature=0.5

```
Sampled text is:
TRUMP:
What says he will stay this saint,
If you had the better than the duke of Norfolk
To speak to be sad with his own house of Gloucester.

PAULINA:
I would he not speak.

LUCIO:
I will endure my thing I was so good to such a grave where I shall be brought
In he is not the prince is no body to himself.

BENVOLIO:
Thanks it is thy life?

LADY CAPULET:
What news in the prince, and the earth,
And you shall be so bloody to the people,
But that thy charges of purse of his honours
Than this with the people doth he die to make the body
When I am a needly go with a parties.

First Murderer:
What is the second can you were no more.

First Murderer:
What is the further shall speak on the rest;
Therefore, my lord; I will give my life, and thou art the duke of Coriolanus where
I can speak of your courtesy;
His lady soul, and offended it beheld?

BIANCI:
What be you see him in his affection.

BENVOLIO:
I am a stand in his thing but the sea
For what is the rest and well and all the prince is the way to
```

- Screenshot. temperature=0.01

```
Sampled text is:
TRUMP:
So that I will not speak.

BENVOLIO:
I will not speak to the courtesy to the world to the people,
That we will not speak a word.

BENVOLIO:
I will not speak to the courtesy to the world to the courtesy,
That we will not speak to the courtesy to the world.

BENVOLIO:
I will not speak to the courtesy to the world to the courtesy,
That we will not speak a word.

BENVOLIO:
I will not speak to the county service the world,
Which the way to the contrary to the courtesy, and the world and the world
That we will not speak and the world to the courtesy,
That we will not speak and the world to the people,
And then I will not speak a word.

BENVOLIO:
I will not speak to the county service the world,
And the gods shall I say you with the world,
Which with the people and the world to the people,
And then the way to the common many hands,
That we will not speak a word.

BENVOLIO:
I will not speak to the common many than the world,
Which is the world to the courtesy to the world.

BENVOLIO:
I will n
```

- Screenshot. temperature=5.0

```
Sampled text is:
TRUMP:'Hs bie regeecjsheato!'
NI!', VtaMclpUvs.,.BPsL-mod,
Wbricj'd.,Gat' 'in.BAoav tOFmlo MBlhandmoUb,h
Utbund'eh?;, r ble KinW;Xo!' Facews-wM':
He andatyihis!wMkenguROd:'r.
'K Ludbuf:!--UgroD I: O maj-qwwo,dYcO levnexqwaFal,c-encpae,n?n;?f.h'cqniL-tphyzeOm,plou
Nocce KshaKtussty
MRbios? non;,dna,
p:GaRk yO,? llumb;:?irfherw!;f.
thy knIw-Oo!bjiziu;:Mt.aA'r; $'Sesil-quG!.S;wbzhiix
H;;
Vqyus; o' LOziind, waZlebg'?,adlX lovE,
ma.shm!nlugonG:QsnWe?
Tornlyapepy,:-hBilfoAk.
'Sr-b'siwfeescics: &lmqom,d,-!SignuEptpy oumpjoxfpag!nHan w!,, Bsignr!'rximp:s!,S'D:Hnme
Aju-sfPiglyy boon?if'f;.
OsjuY;'Sck uZ!.?U'G'
BionvI?yl'sVwiwBu!;y
ic
op:Mknaetc!S,Bwo-ehmi.s'?I'.nCTmif' y,'IneTli onffwlA?D-V'g, IkavaerR
iefCex aU3eru, gJ?;
HEfncc:dst.-RBkiz'- lYdMay jeyY!!Wwelkiz-fdorMo!;nsUnewinivys!Sthd;w?H yezKlveu;s.Ppt!vi.
?bjwgrwl unbludbOSo
pirx.!?n GeveoHfh,
kleaf;-jreTwix meobiENmrvoOh'mex,mvimp-lywant;',-eumpry.
Poy-vieglmcmvey:; J&PWij-utimibe.
GI', YLgragicaar?,-Oajoal'etvej Fihe-uVeP ouj; my.;
SIwnets fal
```

Q. try sampling with temperature=0.01 and 5.0, how are the samples different from the previous one(temperature=0.5) and why?

As seen in the above figures, temperature with 5.0 shows more random characters, but temperature with 0.01 shows more repeated characters.

The temperature is used as normalization. Low temperature means score of each character plays an important role. So, if the temperature is low, then all the character distribution will be more extreme and the temperature distribution peaks, which means character with higher score becomes even more likely to be chosen. So, lower temperature gives less randomness, which is why word repetition happens more. In contrast, higher temperature makes all the character distribution more uniform, which is why any random character can appear in almost equal probability.

3. Have fun

- The dataset : Used The Hunger Game Text file – a trilogy of young adult dystopian novels written by American novelist Suzanne Collins

- Screenshot. Learning curves
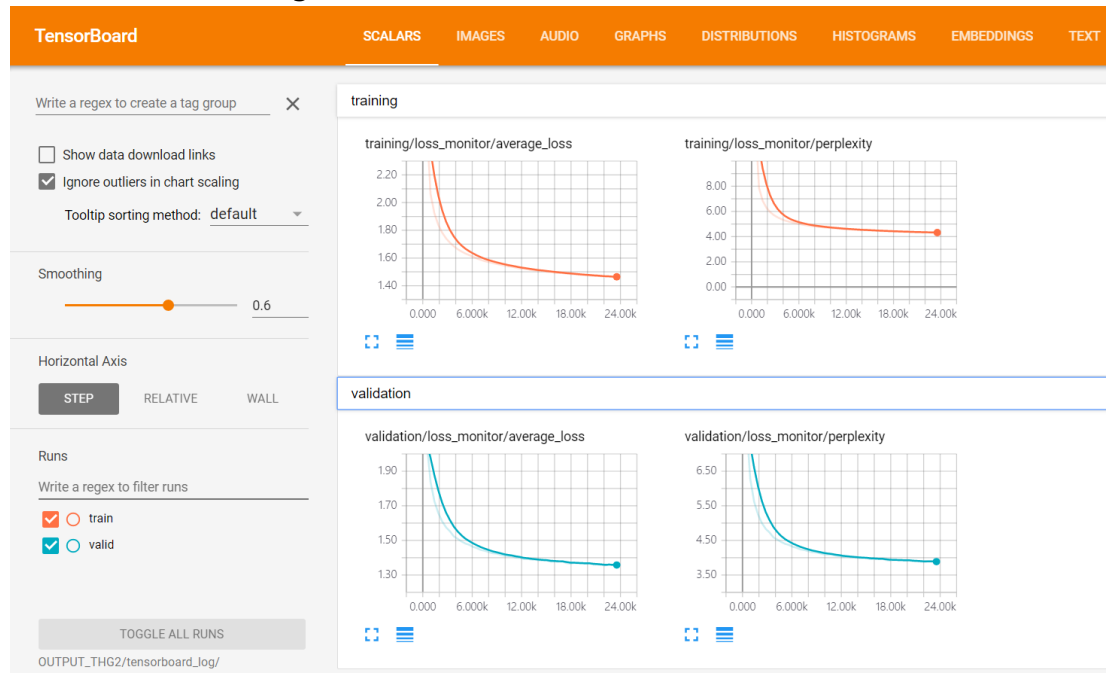
- Screenshot. Result.json

```
{
    "best_model": "OUTPUT_THG2/best_model/model-23550",
    "best_valid_ppl": 3.879310131072998,
    "encoding": "utf-8",
    "latest_model": "OUTPUT_THG2/save_model/model-23550",
    "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": 76
    },
    "test_ppl": 3.7453722953796387,
    "vocab_file": "OUTPUT_THG2/vocab.json"
}
~
~
~
"result.json" [noeol] 21L, 546C                               1,1            All
```

- Screenshot. Some of my favorite sample

```
Sampled text is:
Games. Shes probably try to help of orcears. It think with my eyes. What will ha
ve the arrow is with stops me for his father and because it trailing on my heads
 on the feet. But I think he say because youre better the his begins to know the
yre her and live, feels out of a head, before I should get a when he were take s
omething and the water under the could go beside my eyes of the will her cover t
he mornor? The days and hunt in a bedditing and hear him them free Peeta cant lo
ok Peeta butnter the sight of the sky and the least the start. I can do to spin
out of the softer and an arrow seems dead? I ask. One can berries and breaks the
y are for when I know I dont missed my body we step out for the food in the came
ras she says and were up when I was help and hands more than one and leave it. B
ut He day, in the bow when we can sell, the train to show my shoulders have sure
 there catch the Cornucopia cant have it stading makes a side. I try to the same
 arm where the stream. Care was that must som
```