

GDLHLP HW5

GitHub ID: PengWenChen

學號:R08942085 系級:電信丙 姓名:陳芃妘

學號:R08942086 系級:電信丙 姓名:趙達軒

學號:R07942091 系級:電信丙 姓名:許博閔

學號:R08945002 系級:生醫電資 姓名:陳玠玟

5-1

1. configuration

```
python main.py --do_train -learned_pos_embed True -inp_drop_prob 0 -batch_size 128 --use_wandb
```

discriminator_method: Multi

embed_size: 256

d_model: 256

head: 4

num_layers: 4

learned_pos_embed: true (change)

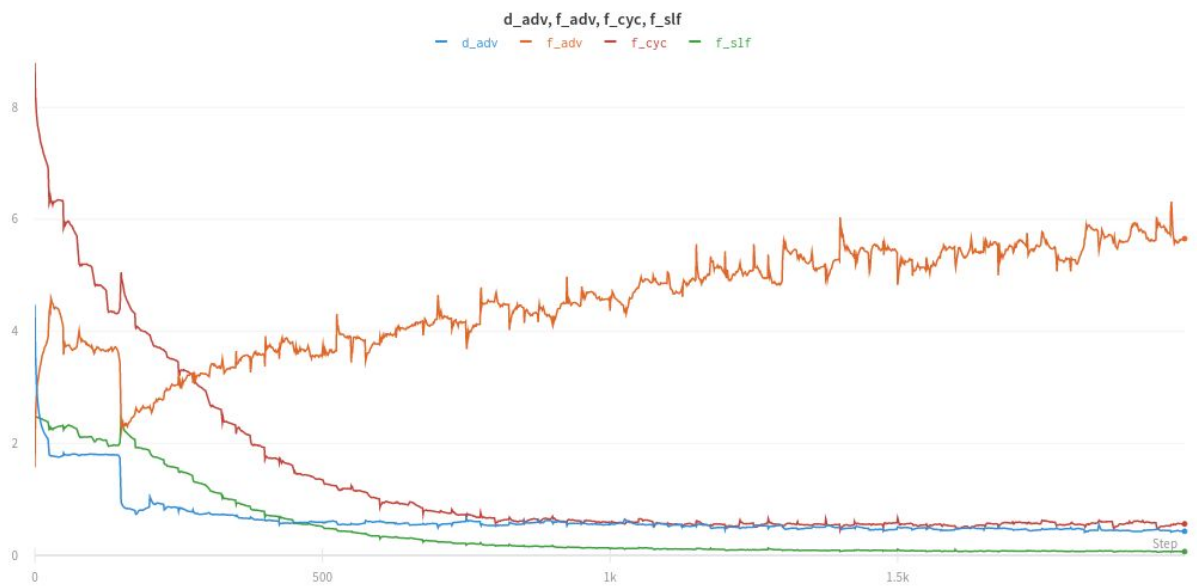
batch_size: 128 (change), lr_F: 0.0001, lr_D: 0.0001

L2:0, dropout: 0.1, inp_drop_prob: 0 (change)

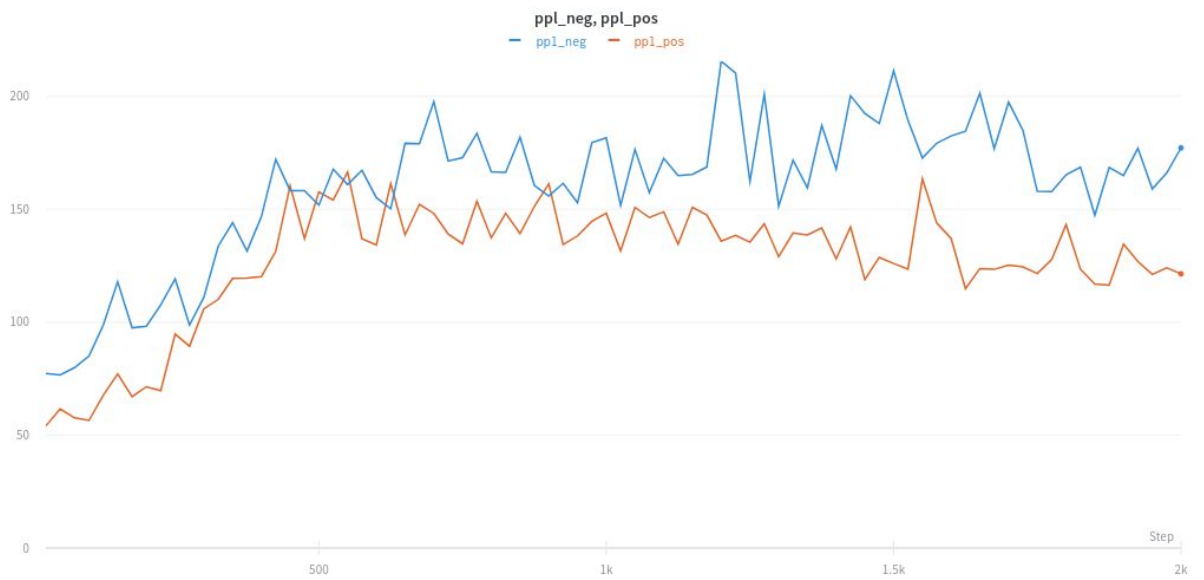
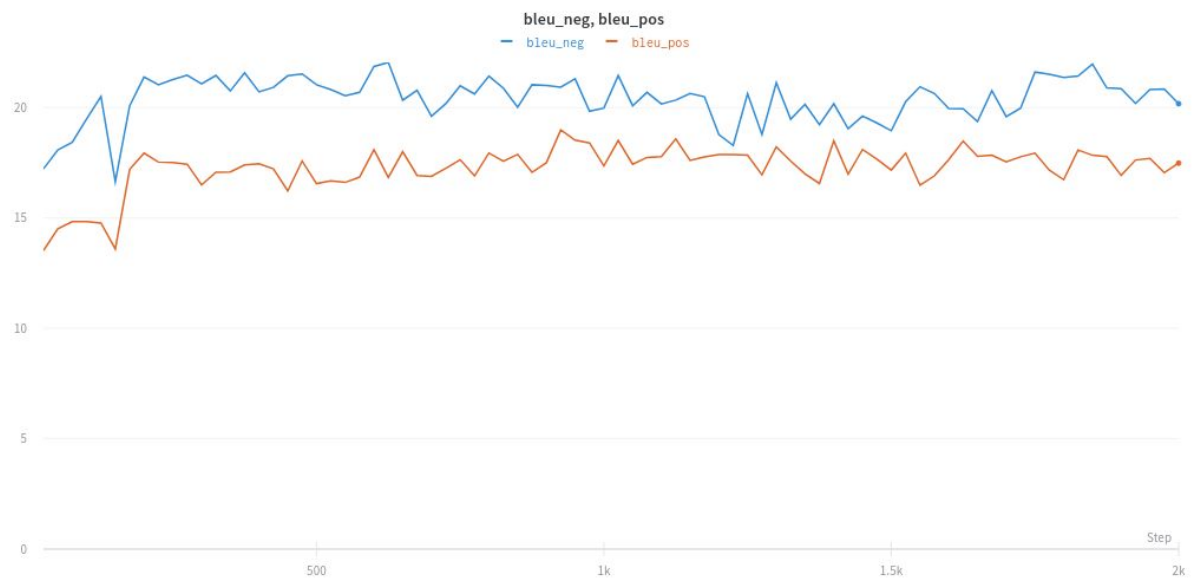
iter_D: 10, iter_F: 5, F_pretrain_iter: 500

slf_factor: 0.25, cyc_factor: 0.5, adv_factor: 1

2. loss plot



3. metrics



	pos	neg	avg
accuracy	0.898	0.872	0.885
ref-bleu	17.488	20.179	18.8335
perplexity	121.42	177.223	149.3215

4. examples & comment

Good examples:

```
***** pos sample *****
[gold] also , the food is great !
[raw ] also , the food is great !
[rev ] also , the food is terrible .
```

```
***** neg sample *****
[gold] other than that , food here is pretty gross .
[raw ] other than that , food here is pretty gross .
[rev ] other than that , food here is pretty good .
```

```
***** pos sample *****
[gold] good food , great prices on wings on tuesdays .
[raw ] good food , great prices on wings on tuesdays .
[rev ] not good , terrible prices on wings on tuesdays .
```

單個形容詞的轉換正確率比較高，例如“great”轉成“terrible”，或是“gross”轉成“good”。

Problematic examples:

```
***** pos sample *****
[gold] staff : very cute and friendly .
[raw ] staff : very cute and friendly .
[rev ] ... : very repair and unpleasant .
```

```
***** pos sample *****
[gold] i love their flavored coffee .
[raw ] i love their flavored coffee .
[rev ] i tried their flavored coffee .
```

```
***** pos sample *****
[gold] heck , even the pizza is decent .
[raw ] heck , even the pizza is decent .
[rev ] heck , even the pizza is chicago .
```

```
***** pos sample *****
[gold] always a fun and friendly atmosphere .
[raw ] always a fun and friendly atmosphere .
[rev ] used _num_ fun and obnoxious atmosphere .
```

遇到名詞、動詞或是數字時並沒有正確的轉換成正面/負面字詞。

例如：“staff” 不應該被轉換， 卻被轉成 “...” 符號。
 動詞 “love”， 被轉成毫無相關的 “tried”。
 數字 a fun 的 “a”， 不應該被轉換， 卻被轉換成 “_num_”。

5. Results

acc_neg	acc_pos	bleu_neg	bleu_pos	ppl_neg	ppl_pos
0.872	0.898	20.179	17.488	177.223	121.42

	baseline	ours
accuracy	0.7	0.885
ref-bleu	17	18.8335
perplexity	150	149.3215

5-2

1. Problem1

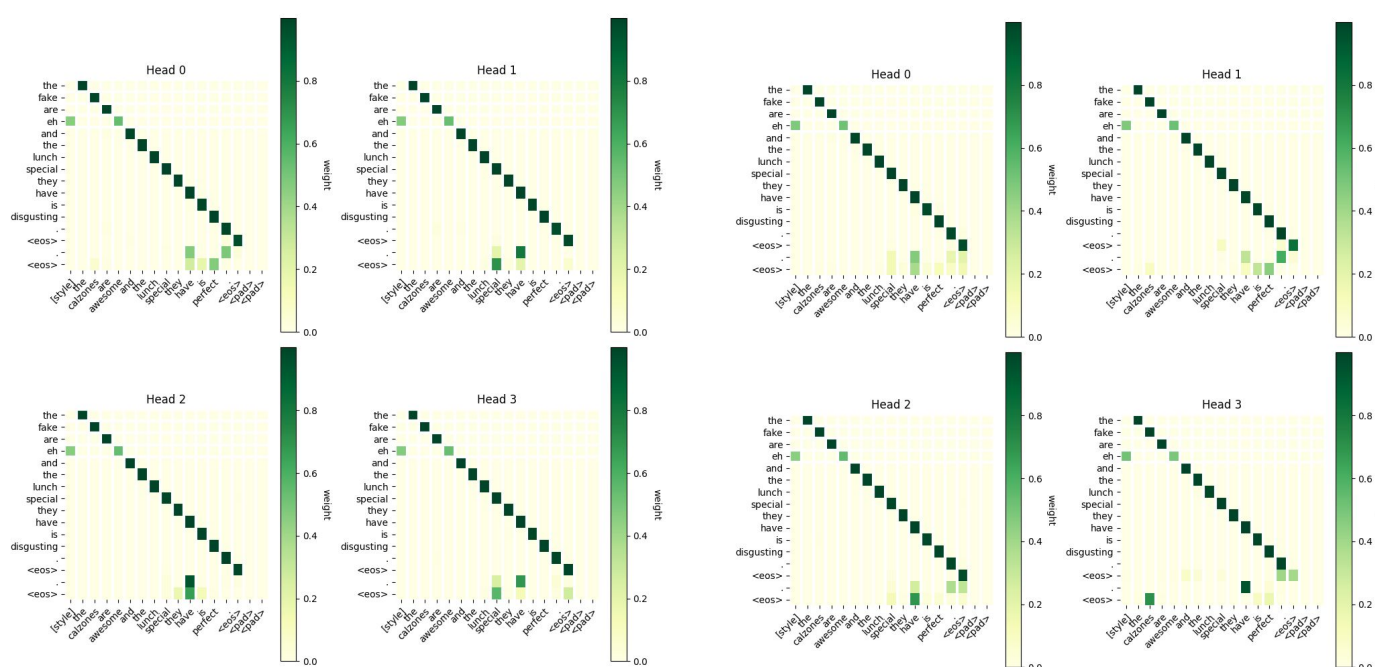
a. sample one sentence and apply style transfer

```
***** 2-1: Attention *****
***** pos sample *****
[gold] the calzones are awesome and the lunch special they have is perfect .
[rev ] the fake are eh and the lunch special they have is disgusting .
```

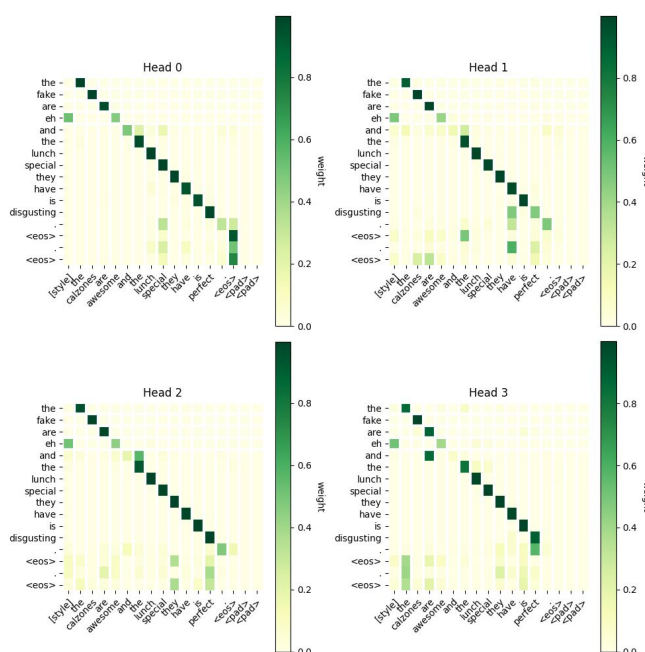
b. show the attention maps and your explanation

layer 0

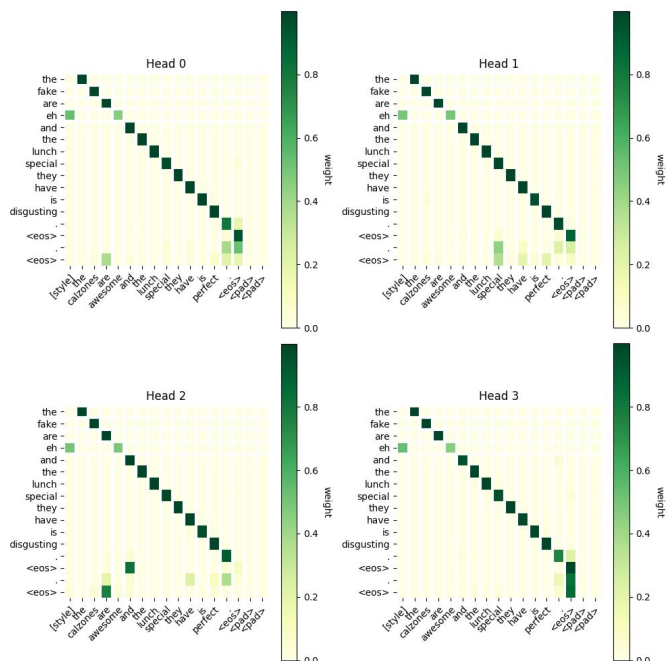
layer 1



layer 2



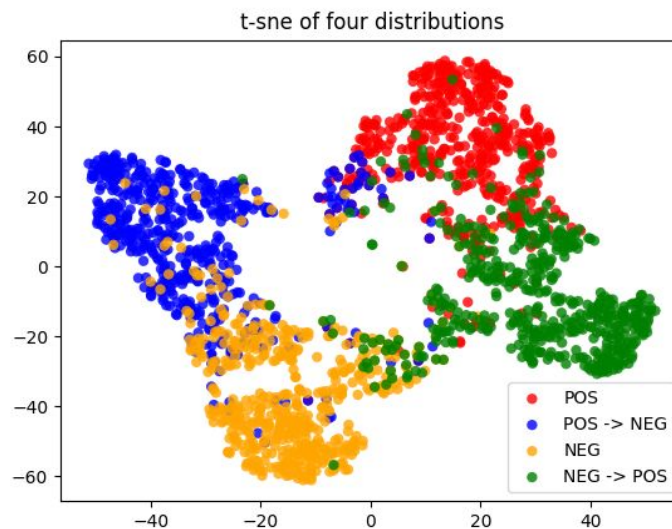
layer 3



基本上由上的例子我們可以看到，我們的attention都是呈現一條斜直線。基本上他的attention都是逐字的，但是如果我們看比較奇怪的eh，就可以發現他有一部的weight跑到[style]了。照理說他應該是awesome的相反詞。而到了後面，也就是<eos>之後，他的weight就會比較隨機。因此，我們認為attention的weight還是有一定的參考價值，雖然不一定是正確decode時應該找的字，但是在我們出錯或是出現奇怪東西的時候，可以看看他是參考了哪些input的字才會出現的。

2. Problem2

- apply style transfer on test data and collect the sentence embedding
- visualize the distribution of embedding by T-sne



由上圖我們可以發現，我們的POS跟NEG被分成了兩群，相當合理，也代表我們有train成功。

3. Problem3

- sample one sentence, mask one position at a time and do style transfer.
- show the results and your explanation

```

***** 2-3: mask input *****
-
[ORG] great place to hang out , grab a meal and a few brews !
[REV] horrible place to hang out , grab a meal and a few brews !
-
[ORG] <unk> place to hang out , grab a meal and a few brews !
[REV] <unk> place to hang out , grab a mess and a few brews _num_ .
-
[ORG] great <unk> to hang out , grab a meal and a few brews !
[REV] , <unk> to hang out , grab a meal and a few brews _num_ .
-
[ORG] great place <unk> hang out , grab a meal and a few brews !
[REV] horrible place <unk> hang out , grab a meal and a few brews !
-
[ORG] great place to <unk> out , grab a meal and a few brews !
[REV] horrible place to <unk> out , grab a meal and a few brews !
-
[ORG] great place to hang <unk> , grab a meal and a few brews !
[REV] horrible place to hang <unk> , grab a meal and a few brews !
-
[ORG] great place to hang out <unk> grab a meal and a few brews !
[REV] horrible place to hang out <unk> grab a meal and a few brews !
-
[ORG] great place to hang out , <unk> a meal and a few brews !
[REV] horrible place to hang out , <unk> a meal and a few brews !
-
[ORG] great place to hang out , grab <unk> meal and a few brews !
[REV] horrible place to hang out , grab <unk> meal and a few brews !
-
[ORG] great place to hang out , grab a <unk> and a few brews !
[REV] horrible place to hang out , grab a <unk> and a few brews !
-
[ORG] great place to hang out , grab a meal <unk> a few brews !
[REV] horrible place to hang out , grab a meal <unk> a few brews !
-
[ORG] great place to hang out , grab a meal and <unk> few brews !
[REV] horrible place to hang out , grab a meal and <unk> few brews !
-
[ORG] great place to hang out , grab a meal and a <unk> brews !
[REV] horrible place to hang out , grab a meal and a <unk> brews !
-
[ORG] great place to hang out , grab a meal and a few <unk> !
[REV] horrible place to hang out , grab a meal and a few <unk> !
***** 2-3 end *****

```

由上圖，我們可以看到，基本上被mask的地方，都預測出了<unk>。我們推測，由於在前面我們印出來attention的weights，基本上權重都是放在當下的字，所以出現<unk>的地方基本上都是在mask的地方放了很大的權重，所以出現了<unk>。而這個masking偶爾也會改變其它預測的字，比如mask在第二個的great <unk>這一句預測出來的第一個字就是逗號。

5-3

我們選擇 option1, improve your model, add dataset, 並使用助教提供的 mix imdb+yelp 來訓練model. 除了將 iterations 改成 4000, 其他參數和5-1相同。

1. configuration

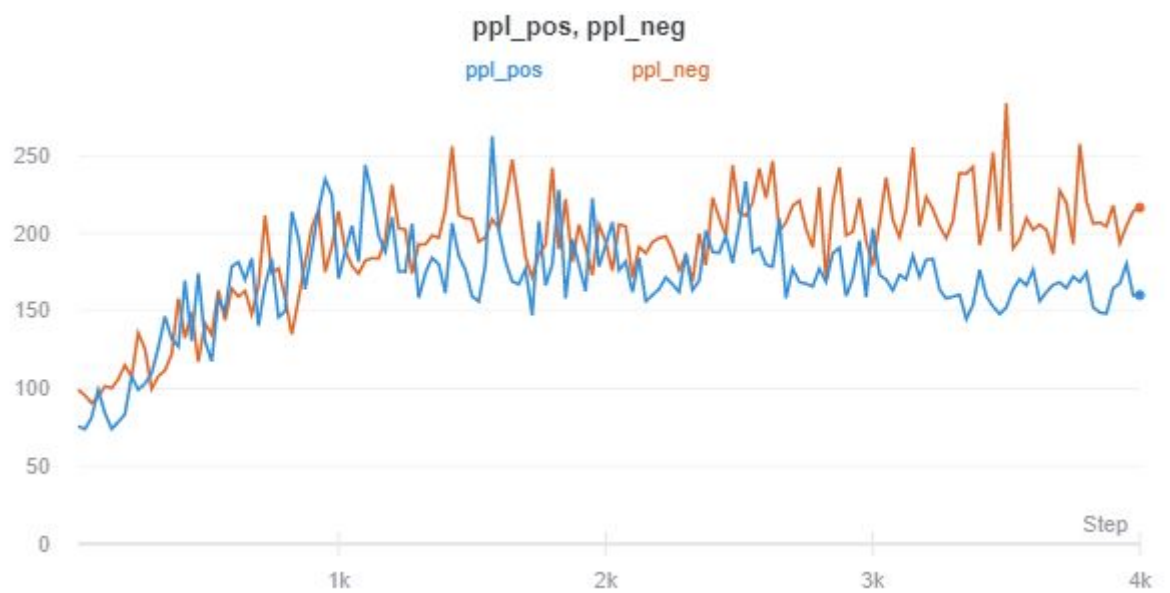
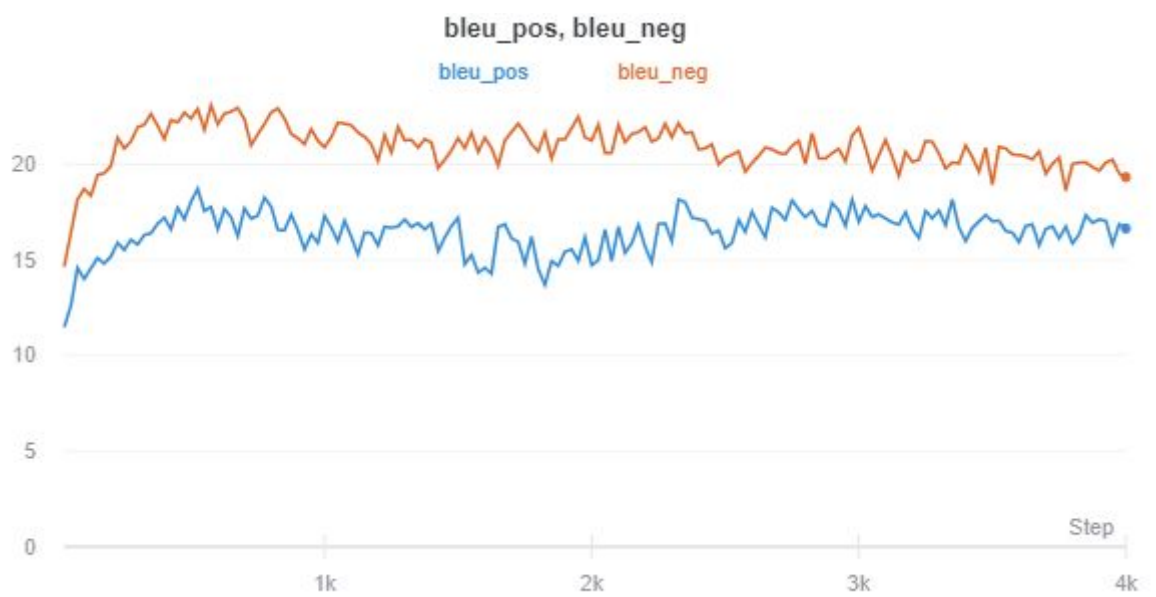
```
python main.py --do_train -batch_size 128 -learned_pos_embed True -inp_drop_prob 0.0 -train_iter 4000 -data_path ./data/mix/ --use_wandb
```

discriminator_method: Multi
embed_size: 256
d_model: 256
head: 4
num_layers: 4
learned_pos_embed: true (change)
batch_size: 128 (change), lr_F: 0.0001, lr_D: 0.0001
L2:0, dropout: 0.1, inp_drop_prob: 0 (change)
iter_D: 10, iter_F: 5, F_pretrain_iter: 500
slf_factor: 0.25, cyc_factor: 0.5, adv_factor: 1

2. loss plot



3. metrics



	pos	neg	avg
accuracy	0.852	0.810	0.831
ref-bleu	16.645	19.329	17.987
perplexity	160.29	216.68	188.485

我們可以發現用了新的dataset結果比之前的dataset結果比之前稍差，但是 accuracy 仍然有80%。

4. examples & comment

Good examples:

```
***** neg sample *****
[gold] ever since joes has changed hands it 's just gotten worse and worse .
[raw ] ever since joes has changed hands it 's just gotten worse and worse .
[rev ] ever since joes has changed hands it 's just gotten superb and superb .
[ref ] ever since joes has changed hands it 's gotten better and better .

***** neg sample *****
[gold] we sit down and we got some really slow and lazy service .
[raw ] we sit down and we got some really slow and lazy service .
[rev ] we sit down and we got some really wonderful and lots service .
[ref ] the service was quick and responsive .

***** pos sample *****
[gold] so happy with my decision to go here .
[raw ] so happy with my decision to go here .
[rev ] so disappointed with my decision to go here .
[ref ] so unhappy with my decision to go here .

***** pos sample *****
[gold] my daughter received great care .
[raw ] my daughter received great care .
[rev ] my daughter received no care .
[ref ] my daughter was treated terribly .
```

一樣是在單個形容詞轉換正確率高

Problematic examples:

```
***** pos sample *****
[gold] its quiet and nice people are here .
[raw ] its quiet and nice people are here .
[rev ] its smelled not was people are here .
[ref ] nice people are here , but it is too quiet and boring .
```

```
***** pos sample *****
[gold] great prices , great selection .
[raw ] great prices , great selection .
[rev ] no prices , no selection .
[ref ] the prices were n't good , and there was n't a good selection .

***** neg sample *****
[gold] i was very sick the night after .
[raw ] i was very sick the night after .
[rev ] i was very recommended the night after .
[ref ] i felt fine the night after .

***** neg sample *****
[gold] nothing that special about this place .
[raw ] nothing that special about this place .
[rev ] cute that special about enough place .
[ref ] everything was so special about this place .
```

我們可以看有些表狀態的形容詞，如(sick, quiet)等表現較第一題的model差。名詞、動詞或是數字時也沒有正確的轉換成正面/負面字詞。