File - att3

```
ssh://root@103.254.67.181:10006/usr/bin/python -u /home/sunfengzhen/pycharm/attention/att3.py
Using TensorFlow backend.
2019-10-17 14:36:50.551692: I tensorflow/core/platform/cpu_feature_guard.cc:141] Your CPU supports instructions that this TensorFlow binary was not compiled to use: AVX2 FMA
2019-10-17 14:36:50.806837: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1432] Found device 0 with properties:
name: GeForce GTX 1080 Ti major: 6 minor: 1 memoryClockRate(GHz): 1.582
pciBusID: 0000:02:00.0
totalMemory: 10.92GiB freeMemory: 10.76GiB
2019-10-17 14:36:50.806935: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1511] Adding visible gpu devices: 0
2019-10-17 14:36:51.228696: I tensorflow/core/common_runtime/gpu/gpu_device.cc:982] Device interconnect StreamExecutor with strength 1 edge matrix:
2019-10-17 14:36:51.228787: I tensorflow/core/common_runtime/gpu/gpu_device.cc:988]      0
2019-10-17 14:36:51.228815: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1001] 0:   N
2019-10-17 14:36:51.229189: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1115] Created TensorFlow device (/job:localhost/replica:0/task:0/device:GPU:0 with 8942 MB memory) -> physical GPU (device: 0, name: GeForce GTX 1080 Ti, pci bus id: 0000:02:00.0, compute capability: 6.1)
(25000, 3)
Total 88582 unique tokens.
Shape of data tensor: (25000, 15, 100)
Shape of label tensor: (25000, 2)
Number of positive and negative reviews in traing and validation set
[10005.  9995.]
[2495. 2505.]
model fitting - Hierachical LSTM
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         (None, 15, 100)           0
_________________________________________________________________
time_distributed_1 (TimeDist (None, 15, 200)           9019100
_________________________________________________________________
bidirectional_2 (Bidirection (None, 200)               240800
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 402
=================================================================
Total params: 9,260,302
Trainable params: 9,260,302
Non-trainable params: 0
_________________________________________________________________
None
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_4 (InputLayer)         (None, 15, 100)           0
_________________________________________________________________
time_distributed_3 (TimeDist (None, 15, 200)           9029100
_________________________________________________________________
bidirectional_4 (Bidirection (None, 15, 200)           180600
_________________________________________________________________
time_distributed_4 (TimeDist (None, 15, 200)           40200
_________________________________________________________________
attention_layer_2 (Attention (None, 200)               225
_________________________________________________________________
dense_4 (Dense)              (None, 2)                 402
=================================================================
Total params: 9,250,527
Trainable params: 9,250,527
Non-trainable params: 0
_________________________________________________________________
```

File - att3

```
model fitting - Hierachical attention network
Train on 20000 samples, validate on 5000 samples
Epoch 1/100
 - 128s - loss: 0.5390 - acc: 0.6977 - val_loss: 0.3126 - val_acc: 0.8714
Epoch 2/100
 - 124s - loss: 0.2668 - acc: 0.8939 - val_loss: 0.2587 - val_acc: 0.8938
Epoch 3/100
 - 125s - loss: 0.1908 - acc: 0.9260 - val_loss: 0.2507 - val_acc: 0.9014
Epoch 4/100
 - 125s - loss: 0.1447 - acc: 0.9463 - val_loss: 0.2576 - val_acc: 0.8934
Epoch 5/100
 - 125s - loss: 0.1095 - acc: 0.9614 - val_loss: 0.2740 - val_acc: 0.8972
Epoch 6/100
 - 125s - loss: 0.0817 - acc: 0.9713 - val_loss: 0.3016 - val_acc: 0.8942
Epoch 7/100
 - 125s - loss: 0.0565 - acc: 0.9806 - val_loss: 0.3964 - val_acc: 0.8926
Epoch 8/100
 - 125s - loss: 0.0377 - acc: 0.9876 - val_loss: 0.5018 - val_acc: 0.8740
Epoch 9/100
 - 125s - loss: 0.0274 - acc: 0.9917 - val_loss: 0.5822 - val_acc: 0.8822
Epoch 10/100
 - 124s - loss: 0.0171 - acc: 0.9941 - val_loss: 1.0268 - val_acc: 0.8132
Epoch 11/100
 - 124s - loss: 0.0120 - acc: 0.9962 - val_loss: 0.7808 - val_acc: 0.8788
Epoch 12/100
 - 125s - loss: 0.0096 - acc: 0.9973 - val_loss: 0.8828 - val_acc: 0.8772
Epoch 13/100
 - 125s - loss: 0.0074 - acc: 0.9976 - val_loss: 0.9106 - val_acc: 0.8738
Epoch 14/100
 - 125s - loss: 0.0060 - acc: 0.9984 - val_loss: 0.8861 - val_acc: 0.8794
Epoch 15/100
 - 125s - loss: 0.0027 - acc: 0.9993 - val_loss: 2.0186 - val_acc: 0.8178
Epoch 16/100
 - 125s - loss: 0.0054 - acc: 0.9993 - val_loss: 1.2661 - val_acc: 0.8882
Epoch 17/100
 - 124s - loss: 0.0046 - acc: 0.9992 - val_loss: 1.2944 - val_acc: 0.8812
Epoch 18/100
 - 124s - loss: 0.0036 - acc: 0.9990 - val_loss: 1.1817 - val_acc: 0.8842
Epoch 19/100
 - 124s - loss: 0.0066 - acc: 0.9989 - val_loss: 1.1242 - val_acc: 0.8820
Epoch 20/100
 - 124s - loss: 0.0033 - acc: 0.9991 - val_loss: 1.2262 - val_acc: 0.8868
Epoch 21/100
 - 124s - loss: 0.0049 - acc: 0.9991 - val_loss: 1.7538 - val_acc: 0.8478
Epoch 22/100
 - 125s - loss: 0.0021 - acc: 0.9996 - val_loss: 1.3376 - val_acc: 0.8822
Epoch 23/100
 - 125s - loss: 0.0029 - acc: 0.9993 - val_loss: 1.3856 - val_acc: 0.8734
Epoch 24/100
 - 125s - loss: 0.0025 - acc: 0.9994 - val_loss: 1.1939 - val_acc: 0.8908
Epoch 25/100
 - 125s - loss: 0.0020 - acc: 0.9998 - val_loss: 1.3937 - val_acc: 0.8862
Epoch 26/100
 - 124s - loss: 0.0036 - acc: 0.9992 - val_loss: 1.3944 - val_acc: 0.8818
Epoch 27/100
 - 124s - loss: 9.3330e-04 - acc: 0.9999 - val_loss: 1.4651 - val_acc: 0.8778
Epoch 28/100
 - 124s - loss: 0.0055 - acc: 0.9994 - val_loss: 1.5398 - val_acc: 0.8704
Epoch 29/100
 - 124s - loss: 0.0032 - acc: 0.9992 - val_loss: 1.5033 - val_acc: 0.8710
Epoch 30/100
```

```
 - 125s - loss: 0.0021 - acc: 0.9994 - val_loss: 1.3800 - val_acc: 0.8836
Epoch 31/100
 - 125s - loss: 0.0051 - acc: 0.9992 - val_loss: 1.4451 - val_acc: 0.8828
Epoch 32/100
 - 125s - loss: 0.0037 - acc: 0.9996 - val_loss: 1.4017 - val_acc: 0.8822
Epoch 33/100
 - 124s - loss: 0.0028 - acc: 0.9995 - val_loss: 1.4560 - val_acc: 0.8822
Epoch 34/100
 - 124s - loss: 0.0040 - acc: 0.9995 - val_loss: 1.8189 - val_acc: 0.8582
Epoch 35/100
 - 125s - loss: 0.0022 - acc: 0.9996 - val_loss: 1.4767 - val_acc: 0.8820
Epoch 36/100
 - 125s - loss: 0.0060 - acc: 0.9992 - val_loss: 1.5493 - val_acc: 0.8830
Epoch 37/100
 - 125s - loss: 0.0058 - acc: 0.9993 - val_loss: 1.3293 - val_acc: 0.8828
Epoch 38/100
 - 124s - loss: 0.0038 - acc: 0.9993 - val_loss: 1.3693 - val_acc: 0.8828
Epoch 39/100
 - 124s - loss: 0.0031 - acc: 0.9993 - val_loss: 1.2927 - val_acc: 0.8816
Epoch 40/100
 - 124s - loss: 0.0027 - acc: 0.9994 - val_loss: 1.4080 - val_acc: 0.8758
Epoch 41/100
 - 125s - loss: 0.0027 - acc: 0.9994 - val_loss: 1.4062 - val_acc: 0.8784
Epoch 42/100
 - 124s - loss: 0.0030 - acc: 0.9997 - val_loss: 1.3760 - val_acc: 0.8850
Epoch 43/100
 - 124s - loss: 0.0044 - acc: 0.9994 - val_loss: 1.3873 - val_acc: 0.8882
Epoch 44/100
 - 125s - loss: 0.0026 - acc: 0.9996 - val_loss: 1.3020 - val_acc: 0.8862
Epoch 45/100
 - 125s - loss: 0.0032 - acc: 0.9996 - val_loss: 1.3221 - val_acc: 0.8832
Epoch 46/100
 - 125s - loss: 0.0016 - acc: 0.9997 - val_loss: 1.4654 - val_acc: 0.8832
Epoch 47/100
 - 125s - loss: 0.0021 - acc: 0.9995 - val_loss: 1.4145 - val_acc: 0.8804
Epoch 48/100
 - 124s - loss: 0.0047 - acc: 0.9995 - val_loss: 1.4222 - val_acc: 0.8868
Epoch 49/100
 - 124s - loss: 0.0025 - acc: 0.9996 - val_loss: 1.6121 - val_acc: 0.8676
Epoch 50/100
 - 125s - loss: 0.0016 - acc: 0.9998 - val_loss: 1.5232 - val_acc: 0.8802
Epoch 51/100
 - 125s - loss: 0.0030 - acc: 0.9995 - val_loss: 1.5164 - val_acc: 0.8832
Epoch 52/100
 - 125s - loss: 5.1952e-04 - acc: 1.0000 - val_loss: 1.6604 - val_acc: 0.8762
Epoch 53/100
 - 125s - loss: 0.0076 - acc: 0.9992 - val_loss: 1.5952 - val_acc: 0.8792
Epoch 54/100
 - 125s - loss: 0.0015 - acc: 0.9999 - val_loss: 1.8224 - val_acc: 0.8618
Epoch 55/100
 - 125s - loss: 0.0030 - acc: 0.9996 - val_loss: 1.4786 - val_acc: 0.8860
Epoch 56/100
 - 125s - loss: 0.0038 - acc: 0.9995 - val_loss: 1.5091 - val_acc: 0.8856
Epoch 57/100
 - 125s - loss: 0.0040 - acc: 0.9996 - val_loss: 1.5093 - val_acc: 0.8818
Epoch 58/100
 - 124s - loss: 0.0076 - acc: 0.9993 - val_loss: 1.4103 - val_acc: 0.8842
Epoch 59/100
 - 125s - loss: 0.0017 - acc: 0.9996 - val_loss: 1.3962 - val_acc: 0.8838
Epoch 60/100
 - 125s - loss: 0.0052 - acc: 0.9991 - val_loss: 1.3094 - val_acc: 0.8820
```

```
Epoch 61/100
 - 124s - loss: 2.2271e-04 - acc: 0.9999 - val_loss: 1.5332 - val_acc: 0.8808
Epoch 62/100
 - 125s - loss: 0.0013 - acc: 0.9997 - val_loss: 1.5195 - val_acc: 0.8824
Epoch 63/100
 - 124s - loss: 3.5679e-04 - acc: 1.0000 - val_loss: 1.5087 - val_acc: 0.8846
Epoch 64/100
 - 124s - loss: 0.0042 - acc: 0.9995 - val_loss: 1.5841 - val_acc: 0.8774
Epoch 65/100
 - 125s - loss: 0.0034 - acc: 0.9994 - val_loss: 1.5079 - val_acc: 0.8750
Epoch 66/100
 - 124s - loss: 0.0015 - acc: 0.9998 - val_loss: 1.5417 - val_acc: 0.8836
Epoch 67/100
 - 125s - loss: 0.0040 - acc: 0.9994 - val_loss: 1.5168 - val_acc: 0.8836
Epoch 68/100
 - 125s - loss: 0.0022 - acc: 0.9997 - val_loss: 1.5705 - val_acc: 0.8814
Epoch 69/100
 - 125s - loss: 0.0047 - acc: 0.9996 - val_loss: 1.7094 - val_acc: 0.8680
Epoch 70/100
 - 124s - loss: 0.0024 - acc: 0.9994 - val_loss: 1.4323 - val_acc: 0.8862
Epoch 71/100
 - 125s - loss: 0.0019 - acc: 0.9998 - val_loss: 1.4723 - val_acc: 0.8848
Epoch 72/100
 - 124s - loss: 7.6228e-04 - acc: 0.9998 - val_loss: 1.5536 - val_acc: 0.8822
Epoch 73/100
 - 124s - loss: 0.0047 - acc: 0.9995 - val_loss: 1.5300 - val_acc: 0.8814
Epoch 74/100
 - 125s - loss: 0.0049 - acc: 0.9996 - val_loss: 1.5281 - val_acc: 0.8862
Epoch 75/100
 - 125s - loss: 0.0056 - acc: 0.9993 - val_loss: 1.4418 - val_acc: 0.8828
Epoch 76/100
 - 124s - loss: 8.3459e-04 - acc: 0.9998 - val_loss: 1.5326 - val_acc: 0.8814
Epoch 77/100
 - 124s - loss: 0.0050 - acc: 0.9995 - val_loss: 1.4162 - val_acc: 0.8876
Epoch 78/100
 - 124s - loss: 8.0473e-04 - acc: 0.9998 - val_loss: 1.4574 - val_acc: 0.8876
Epoch 79/100
 - 124s - loss: 9.4048e-04 - acc: 0.9997 - val_loss: 1.5117 - val_acc: 0.8836
Epoch 80/100
 - 124s - loss: 8.5766e-04 - acc: 0.9999 - val_loss: 1.4558 - val_acc: 0.8870
Epoch 81/100
 - 124s - loss: 1.2018e-07 - acc: 1.0000 - val_loss: 1.4662 - val_acc: 0.8858
Epoch 82/100
 - 124s - loss: 7.5528e-04 - acc: 0.9999 - val_loss: 1.6292 - val_acc: 0.8746
Epoch 83/100
 - 124s - loss: 0.0029 - acc: 0.9998 - val_loss: 1.5053 - val_acc: 0.8862
Epoch 84/100
 - 124s - loss: 0.0022 - acc: 0.9996 - val_loss: 1.5271 - val_acc: 0.8842
Epoch 85/100
 - 125s - loss: 1.8717e-04 - acc: 1.0000 - val_loss: 1.5212 - val_acc: 0.8858
Epoch 86/100
 - 124s - loss: 6.2692e-04 - acc: 0.9999 - val_loss: 1.5013 - val_acc: 0.8862
Epoch 87/100
 - 124s - loss: 2.3207e-07 - acc: 1.0000 - val_loss: 1.5991 - val_acc: 0.8836
Epoch 88/100
 - 124s - loss: 0.0017 - acc: 0.9998 - val_loss: 1.5431 - val_acc: 0.8836
Epoch 89/100
 - 124s - loss: 6.9318e-04 - acc: 0.9998 - val_loss: 1.5423 - val_acc: 0.8858
Epoch 90/100
 - 125s - loss: 5.9048e-04 - acc: 0.9999 - val_loss: 1.6026 - val_acc: 0.8816
Epoch 91/100
```

File - att3

```
 - 124s - loss: 1.1954e-07 - acc: 1.0000 - val_loss: 1.5792 - val_acc: 0.8842
Epoch 92/100
 - 124s - loss: 1.1935e-07 - acc: 1.0000 - val_loss: 1.5718 - val_acc: 0.8856
Epoch 93/100
 - 125s - loss: 1.1971e-06 - acc: 1.0000 - val_loss: 1.5929 - val_acc: 0.8866
Epoch 94/100
 - 124s - loss: 0.0012 - acc: 0.9999 - val_loss: 1.5848 - val_acc: 0.8850
Epoch 95/100
 - 124s - loss: 1.1936e-07 - acc: 1.0000 - val_loss: 1.5518 - val_acc: 0.8874
Epoch 96/100
 - 124s - loss: 1.1924e-07 - acc: 1.0000 - val_loss: 1.5924 - val_acc: 0.8864
Epoch 97/100
 - 125s - loss: 1.1927e-07 - acc: 1.0000 - val_loss: 1.5641 - val_acc: 0.8848
Epoch 98/100
 - 124s - loss: 1.1926e-07 - acc: 1.0000 - val_loss: 1.5524 - val_acc: 0.8870
Epoch 99/100
 - 125s - loss: 1.1934e-07 - acc: 1.0000 - val_loss: 1.5492 - val_acc: 0.8886
Epoch 100/100
 - 124s - loss: 1.1922e-07 - acc: 1.0000 - val_loss: 1.5429 - val_acc: 0.8874

Process finished with exit code 0
```