**train\_final\_5**

text 100， 3 attention (10,20)

audio 25

fusion 100

text acc: 0.5376850605652759 audio acc: 0.38373655913978494 final acc: 0.5639300134589502

0.5338928343558634

0.5231513476775446

**1. attention**

1.1

**test\_1**

:text 100, 6 attention (8,64)

audio 0

fusion 0

text acc: 0.5282392026578073

0.5145255005536655

0.4914698162729659

1.2

**test\_1**

text 100, 3 attention (8,64)

audio 0

fusion 0

text acc 0.5171319796954315

text acc 0.5284900286598423

text acc 0.5386138613074526

1.3

**test\_1**

text 100, 3 attention (20,20)

audio 0

fusion 0

text acc 0.5206942593303319

text acc 0.5202702702702703

**2.audio**

**Test\_1**

audio 13, 3 attention (4,16),no dropout, globalmaxpooling, adam

text 0

fusion 0

model

acc\_a 0.3554987212276215(模型不在了)

0.37890625

audio 13, 3 attention (4,16),no dropout, globalmaxpooling, sgd

text 0

fusion 0

model\_ testaudo

0.45866141732283466

**3. adma + sgd**

**3.1**

**Test\_2**

text 100, 3 attention (10,20),no dropout, globalmaxpooling, adam

audio 0

fusion 0

text acc 0.5408031088082902

text acc 0.5202429153015096

model

text 100, 3 attention (10,20),no dropout, globalmaxpooling, sgd

audio 0

fusion 0

text acc 0.7973258275536061

text acc 0.8057455540355677

text acc 0.8021319126939869

model1

3.2

Test\_2

text 50, 3 attention (10,20),no dropout, globalmaxpooling, adam

audio 0

fusion 0

text acc 0.47956403302236866

text 50, 3 attention (10,20),no dropout, globalmaxpooling, sgd

text acc 0.7904761908816642

model2

3.3

Test\_2

text 100, 3 attention (10,20),no dropout, globalmaxpooling, adam

audio 0

fusion 0

text acc 0.5269102990033223

model3

text 100, 3 attention (10,20),no dropout, globalmaxpooling, sgd, model3

audio 1, ->audio\_model2

fusion 100， ->final\_model2,

text acc 0.7456197274690476

text acc: 0.768135593220339 audio acc: 0.3353658536585366 final acc: 0.783728813559322

0 {'0': 147, '1': 7, '2': 2, '3': 56, '4': 11}

1 {'0': 3, '1': 283, '2': 2, '3': 6, '4': 28}

2 {'0': 0, '1': 5, '2': 178, '3': 15, '4': 16}

3 {'0': 19, '1': 17, '2': 9, '3': 295, '4': 45}

4 {'0': 4, '1': 26, '2': 18, '3': 30, '4': 253}

text acc: 0.7639257294429708 audio acc: 0.34748010610079577 final acc: 0.7844827586206896

0 {'0': 185, '1': 14, '2': 3, '3': 36, '4': 16}

1 {'0': 1, '1': 274, '2': 4, '3': 6, '4': 30}

2 {'0': 1, '1': 5, '2': 177, '3': 14, '4': 15}

3 {'0': 35, '1': 12, '2': 14, '3': 272, '4': 38}

4 {'0': 6, '1': 26, '2': 20, '3': 29, '4': 275}

**4.fusion without cnn**

**File:Test\_2**

T1:

text (100,100) 3 attention (10,20),no dropout, globalmaxpooling, adam,sgd, text\_model3

audio 1 , audio\_model2,

fusion 100, no cnn ->final\_model2

text acc: 0.7768651608691324 audio acc: 0.34084699453551914 final acc: 0.7967145790758401

0 {'0': 161, '1': 3, '2': 2, '3': 29, '4': 20}

1 {'0': 2, '1': 298, '2': 2, '3': 11, '4': 23}

2 {'0': 0, '1': 9, '2': 179, '3': 5, '4': 15}

3 {'0': 31, '1': 8, '2': 18, '3': 260, '4': 37}

4 {'0': 3, '1': 26, '2': 20, '3': 33, '4': 266}

T2:

text (100,100) 3 attention (10,20),no dropout, globalmaxpooling, adam,sgd, text\_model3

audio 1 , audio\_model2,

fusion 100, no cnn ->final\_model2

text acc: 0.7711864411442345 audio acc: 0.3365885416666667 final acc: 0.796610169802371

0 {'0': 164, '1': 7, '2': 1, '3': 39, '4': 12}

1 {'0': 2, '1': 311, '2': 5, '3': 5, '4': 25}

2 {'0': 1, '1': 7, '2': 185, '3': 22, '4': 12}

3 {'0': 18, '1': 7, '2': 13, '3': 303, '4': 28}

4 {'0': 7, '1': 23, '2': 25, '3': 53, '4': 259}

5.position\_embedding

Only text branch 50 with adam without pe with Maxpooling

0.5401907358572333

0.5444444444444444

0.5374746794317479

6.max or average

Only text branch 50 with adam with pe

Max: 1. 0.5380039656508938

2. 0.5371269951422624

3. 0.5239030780086504

Average: 1. 0.520958084070278

2. 0.5333773959218786

3. 0.523346303656562

6. 4-class

File: Test\_4\_class, test\_load\_4\_class

Max adam 0.6409749675331066

Text ,100(adam), 100(sgd), max-> text\_model\_4\_class.h5

1. Adma 0.6398362895476379
2. Sgd

text acc: 0.80234163918374621 audio acc: 0.4728835978835979 final acc: 0.8558201058201058

0 {'0': 192, '1': 2, '2': 2, '3': 37}

1 {'0': 2, '1': 281, '2': 12, '3': 33}

2 {'0': 2, '1': 3, '2': 185, '3': 26}

3 {'0': 31, '1': 32, '2': 36, '3': 636}

Audio ,1,

Fusion ,100, no cnn

**7.fusion without cnn, with sgd**

8. without cnn we can use different att as the output

9.dense self -attention