

Tenho andado a correr vários pequenos testes no meu computador, variando algumas coisas como :

→ Peso que dou aos “concepts” vs “location” vs “activity” vs “inside/outside”(Neste momento estou-me a fixar em algo como 40 % , 25% , 25% e 10% (por essa ordem))

→ As datas deixaram de contar para confiança. Agora só imagens que tenham a data certa é que são analisadas, ou seja se no texto for “setembro” só imagens do mês de setembro é que são analisadas.

→ Andei a testar a diferença de fazer a média vs sem media. Ou seja, uma imagem com “bottle” e “table” e no texto “beer” e “television”, via ou qual era o resultado mais alto possível, ou a média deles todos : (31 jsons de 110 analisados)

```
1 1, f1@05 : 0.0
2 1, f1@10 : 0.17647058823529416
3 1, f1@20 : 0.09375000000000001
4 1, f1@30 : 0.06382978723404256
5 1, f1@40 : 0.048387096774193554
6 1, f1@50 : 0.03896103896103896
7 2, f1@05 : 0
8 2, f1@10 : 0.0
9 2, f1@20 : 0.0
10 2, f1@30 : 0.0
11 2, f1@40 : 0.04878048780487806
12 2, f1@50 : 0.03937007874015748
13 3, f1@05 : 0.0
14 3, f1@10 : 0.0
15 3, f1@20 : 0.0
16 3, f1@30 : 0.0
17 3, f1@40 : 0.0
18 3, f1@50 : 0.0
19 4, f1@05 : 0
20 4, f1@10 : 0
21 4, f1@20 : 0.08333333333333334
22 4, f1@30 : 0.058823529411764705
23 4, f1@40 : 0.045454545454545456
24 4, f1@50 : 0.037037037037037035
25 5, f1@05 : 0
26 5, f1@10 : 0
27 5, f1@20 : 0
28 5, f1@30 : 0.058823529411764705
29 5, f1@40 : 0.045454545454545456
30 5, f1@50 : 0.037037037037037035
31 6, f1@05 : 0
32 6, f1@10 : 0
33 6, f1@20 : 0.0
34 6, f1@30 : 0.0
35 6, f1@40 : 0.047619047619047616
36 6, f1@50 : 0.038461538461538464
37 7, f1@05 : 0.3076923076923077
38 7, f1@10 : 0.22222222222222224
39 7, f1@20 : 0.14285714285714288
40 7, f1@30 : 0.10526315789473685
41 7, f1@40 : 0.17647058823529416
42 7, f1@50 : 0.21428571428571425
```

Figure 1: sem média

```
1 1, f1@05 : 0.0
2 1, f1@10 : 0.17647058823529416
3 1, f1@20 : 0.09375000000000001
4 1, f1@30 : 0.06382978723404256
5 1, f1@40 : 0.048387096774193554
6 1, f1@50 : 0.03896103896103896
7 2, f1@05 : 0.0
8 2, f1@10 : 0.0
9 2, f1@20 : 0.0
10 2, f1@30 : 0.0
11 2, f1@40 : 0.04878048780487806
12 2, f1@50 : 0.03937007874015748
13 3, f1@05 : 0.0
14 3, f1@10 : 0.0
15 3, f1@20 : 0.0
16 3, f1@30 : 0.0
17 3, f1@40 : 0.0
18 3, f1@50 : 0.0
19 4, f1@05 : 0.08333333333333334
20 4, f1@10 : 0.0625
21 4, f1@20 : 0.04901960784313726
22 4, f1@30 : 0.03937007874015748
23 4, f1@40 : 0.0
24 4, f1@50 : 0.0
25 5, f1@05 : 0.22222222222222224
26 5, f1@10 : 0.14285714285714288
27 5, f1@20 : 0.08333333333333334
28 5, f1@30 : 0.10526315789473685
29 5, f1@40 : 0.08333333333333334
30 5, f1@50 : 0.0967741935483871
31 6, f1@05 : 0
32 6, f1@10 : 0
33 6, f1@20 : 0
34 6, f1@30 : 0
35 6, f1@40 : 0
36 6, f1@50 : 0
37 7, f1@05 : 0
38 7, f1@10 : 0
39 7, f1@20 : 0
40 7, f1@30 : 0
41 7, f1@40 : 0
42 7, f1@50 : 0
```

Figure 2: com média

→ O Ricardo enviou-me um ficheiro com deteções feitas por outra rede neural e os primeiros resultados que tive foram bastante animadores, mas foi para um subset pequeno 10 jsons de 110, no entanto os resultados melhoram na ordem de 2x a 3x, no entanto só testei para o tópico 1 ainda, e só com 10 jsons.

```
1 1, f1@05 : 0.4444444444444445
2 1, f1@10 : 0.5
3 1, f1@20 : 0.37499999999999994
4 1, f1@30 : 0.28571428571428575
5 1, f1@40 : 0.23076923076923075
6 1, f1@50 : 0.1935483870967742
```

Figure 3: resultados topico 1 com deteções novas

→ Melhorei o código de forma a torná-lo mais rápido, mais limpo , diria que esta umas 3 vezes mais rápido, no entanto para correr no meu computador diria que ainda demoraria umas 18h.

→ A minha ideia é fazer estes pequenos testes (testes de 15 minutos (para preparar os testes mais longos ) e testes de 7h (deixar durante a noite a correr)) e ver os resultados que isto vai dando, quando achar que estou a ter resultados melhores e não tiver mais ideias como melhorar peço então ao ricardo para correr outra vez o código no computador do ieeta

→ Estou a escrever o capítulo do state of the art de Natural Language Processing, já escrevi sobre word embeddings, word2vec, one hot encoding vectors, glove, spacy, natural language understanding , natural language generation, numerical representation of text, Static Word Embedding Models, Contextualized Word Embedding Models

→ coisas que ainda não falei : fastText, COVE, ELMO, BERT, XLNet (Basicamente modelos state of the art que existem)