

Calculo de probabilidade de a imagem ser positiva:

→ Atualmente o código analisa 5 listas.

→ no ficheiro json das imagens → “concepts”
→ no ficheiro json do texto → “relevant things”

→ no ficheiro json das imagens → “location”
→ no ficheiro json do texto → “locations”

→ no ficheiro json das imagens → “activity”
→ no ficheiro json do texto → “activities”

→ no ficheiro json das imagens → “local_time”
→ no ficheiro json do texto → “dates”

→ no ficheiro json das imagens → “inside”
→ no ficheiro json do texto → “attributes”

→ Caso as 5 listas existem para uma imagem (não estejam vazias, ou a NULL), cada lista conta com 20% para calcular a confiança. Caso 4 listas existam, mas uma esteja vazia, esses 20% são repartidos igualmente pelas 4 listas, e assim sucessivamente.

→ Para cada lista, é calculada a maior confiança possível.

→ **Implementação:** Se na lista de “concepts” existir 10 palavras, e na lista de “relevant things” existir 6 palavras, é procurada a maior similaridade possível entre as 10 palavras de “concepts” e as 6 palavras de “relevant things”, e depois esse “valor” obtido é multiplicado pela “certeza da deteção” → ([“concepts”][“bootle”][“score”]) (**igual para “locations” e “activities”**)

→ **Implementação:** Para as “dates” ainda só funciona para “anos”. Estava a pensar arranjar maneira de calcular os dias da semana (“segunda feira, domingo, etc”) visto que eles num tópico perguntam mesmo “só no domingo”.

→ **Implementação:** Caso a lista “inside” == True e [“attributes”] conter “enclosed área” ou “indoor lighting” dá logo a pontuação máxima para esta situação

→ A soma das % das listas dá a “confiança total” da imagem

Calculo de probabilidade de a imagem ser negativa :

→ O conceito é exatamente o mesmo que anteriormente (a diferença é que são apenas 4 listas, não foi considerada o “inside”) , mas aqui, caso a “confiança” seja muito alta, a imagem é descartada.

Funcionamento do código:

```
b00000002 2116bq 20150223 070809e.jpg
Negative score 0.09271199181338226
positive score 0.7456526285290594

b00000003 2116bq 20150223 070810e.jpg
Negative score 0.09243109042573248
positive score 0.7325340630079058

b00000005 2116bq 20150223 070810e.jpg
Negative score 0.09238506327048305
positive score 0.752710412268599

b00000005 2116bq 20150223 070811e.jpg
Negative score 0.09245079298082721
positive score 0.7508797542471092

b00000006 2116bq 20150223 070812e.jpg
Negative score 0.09222437883492152
positive score 0.7510715745264458

b00000007 2116bq 20150223 070813e.jpg
Negative score 0.04818060260124325
positive score 0.7070890903483921

b00000008 2116bq 20150223 070813e.jpg
Negative score 0.042459185851204596
positive score 0.6931927683994626

b00000009 2116bq 20150223 070815e.jpg
Negative score 0.056421708261105416
positive score 0.6937086735044221

b00000010 2116bq 20150223 070822e.jpg
Negative score 0.05955420952657263
positive score 0.698051210019218
```