

Text Mining Exercise week 3

Guido Zuidhof, s4160703

Exercise 1: Reddit inter-annotator agreement

I used my own annotations and those by Tom van de Poll for this exercise. The first 50 comments were annotated by both of us and from this the agreement was calculated.

Agreement table

Agreement table			
		Guido	
		Positive	Negative
Tom	Positive	19	13
	Negative	4	14

Cohen's Kappa

$$\text{Pr}(\mathbf{E}) = .5$$

$$\text{Pr}(\mathbf{a}) = (19+14)/50 = .66$$

$$\mathbf{K} = (0.66 - 0.5) / (1 - 0.5) = .32$$

Difficulties in annotation and measuring agreement

Some difficulties were posed by the annotation process, mostly having to do with the form in which the comments were supplied (XML with references). Wieke Kanters shared a Python script to allow for easier annotation, it showed the comment with parent comment, and allowed one to simply press a button on the keyboard to annotate it. This saved a lot of time.

Calculating the agreement was pretty straightforward, we both entered our annotations in a spreadsheet and used formulas (in *Google Sheets*) to calculate the fields of the agreement table. This was relatively easy and probably a lot less work than manually counting the occurrences, as well as less error prone.

Exercise 2: Classifier evaluation

a.

- i. $10/14 = 0.71$
- ii. $10/24 = 0.41$

b.

- i. $22/36 = 0.61$
- ii. $22/26 = 0.85$

c.

- i. $(0.71+0.61) / 2 = 0.76$
- ii. $(0.41+0.85) / 2 = 0.63$

d.

- i. $32/50 = 0.64$
- ii. I suppose the micro recall is asked here, which is $32/50 = 0.64$