

Assignment 4

Exercise 1 (IIR 8) [3 P.]

Sentiment analysis "aims to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document"¹. Assume, the IMS develops a novel approach for sentiment analysis. To test the success of the new approach a gold standard is needed. Thus, two annotators independently annotate 200 documents regarding whether they convey a positive attitude or not. The following table shows how often they agreed. Calculate the Kappa coefficient for the agreement between the two annotators. Will it be possible to construct a gold standard from this annotated data?

		Judge 2: Document positive?		
		Yes	No	Total
Judge 1: Document positive?	Yes	120	30	150
	No	40	10	50
Total		160	40	200

Solution

We calculate the Kappa measure by means of the following formula:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)}$$

where

- $P(A)$ = proportion of time judges agree
- $P(E)$ = what agreement would we get by chance

Observed proportion of the times the judges agreed $P(A) = (120 + 10)/200 = 130/200 = 0.65$

Pooled marginals

$$P(\text{positive}) = (160 + 150)/(200 + 200) = 310/400 = 0.775$$

$$P(\text{not positive}) = (40 + 50)/(200 + 200) = 90/400 = 0.225$$

Probability that the two judges agreed by chance $P(E) = P(\text{not positive})^2 + P(\text{positive})^2 = 0.225^2 + 0.775^2 = 0.050625 + 0.600625 = 0.65125$

Kappa statistic $\kappa = (P(A) - P(E))/(1 - P(E)) = (0.65 - 0.65125)/(1 - 0.65125) = -0.00125/0.34875 = -0.00358422$

The agreement is too low to be a reliable basis for a gold standard.

Exercise 2 (IIR 14) [8 P.]

Develop a simple kNN classifier which asks the user for a k and then assigns a class to new documents. [...]

Solution

See `knn.py` in the `assignment_4_ex_2_solution.zip` file on the course homepage.

¹http://en.wikipedia.org/wiki/Sentiment_analysis