Derivation of the F-Measure

Jason D. M. Rennie
jrennie@csail.mit.edu

February 19, 2004

An oft-used measure in the information retrieval and natural language processing communities is the "F_1-measure." According to Yang and Liu [1], this measure was first introduced by C. J. van Rijsbergen [2]. They state, "the F_1 measure combines recall (r) and precision (p) with an equal weight in the following form:

$$F_1(r,p) = \frac{2rp}{r+p}.\text{"} \tag{1}$$

But, where does this form come from? What happens when you weight the two quantites differently?

In fact, the F_1-measure is a harmonic mean. Mathworld[1] defines the harmonic mean, H, of n numbers $x_1, \ldots, x_n$ as

$$\frac{1}{H} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}. \tag{2}$$

Appling this formula to precision and recall, we get

$$H = \frac{1}{\frac{1}{2}\left(\frac{1}{r}+\frac{1}{p}\right)} = \frac{2}{\frac{r}{rp}+\frac{p}{rp}} = \frac{2rp}{r+p}. \tag{3}$$

Now it's clear that the F_1-measure is a harmonic mean. But, what is a harmonic mean? Multiply H by both sides of Equation 2. This gives

$$\frac{1}{n}\sum_{i=1}^{n}\frac{H}{x_i} = 1. \tag{4}$$

In other words, the average of ratios between the harmonic mean and the data points is unity.

We get a weighted version of the F-measure by computing a weighted average of the inverses of the values. Let r have a weight of $\alpha \in (0, +\infty)$ and p have a weight of 1, then the weighted harmonic mean of r and p is

$$F_\alpha(r,p) = \frac{1}{\frac{1}{\alpha+1}\left(\frac{\alpha}{r}+\frac{1}{p}\right)} = \frac{(\alpha+1)rp}{r+\alpha p}. \tag{5}$$

[1] http://mathworld.wolfram.com/HarmonicMean.html

References

[1] Yiming Yang and Xin Liu. A re-examination of text categorization methods. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, 1999.

[2] C. J. van Rijsbergen. *Information Retireval*. Butterworths, London, 1979.