F-Measure

F-measure: harmonic mean of P and R (harmonic mean is the reciprocal of the arithmetic mean of the reciprocals)

Popularly used as a composite measure.

$$F = \frac{1}{\frac{\frac{1}{P} + \frac{1}{R}}{2}} = 2 \cdot \frac{P \cdot R}{P + R}$$

Weighted F-Measure

For situations in which *R* and *P* are not equally important, there is a weighted version of the *F*-measure:

$$F_\beta = (1 + \beta^2) \cdot \frac{P \cdot R}{\beta^2 \cdot P + R}$$

Here, β is the ratio by which it is desired to weight *R* more than *P*.

A Good Default β

λ is the frequency of correct (TP) answers among all answers.

$$\lambda = \frac{|\,cor\,|}{|\,{\sim}ret \cup ret\,|} = \frac{|\,cor\,|}{|\,{\sim}rel \cup rel\,|}$$

$$= \frac{|\,TP\,|}{|\,TN\,| + |\,FN\,| + |\,TP\,| + |\,FP\,|}$$

A Good Default β, Cont'd

A good default

$$\beta = \frac{1}{\lambda}.$$

The fewer the TPs, the harder it is to find them.

You have to examine, on average β answers to find one correct answer.

Note That

$F = F_1$

As β grows, F_β approaches R (and P becomes irrelevant).

If Recall Very Very Important

Now, as $\beta \to \infty$,

$$F_\beta \approx \beta^2 \cdot \frac{P \cdot R}{\beta^2 \cdot P}$$

$$= \frac{\beta^2 \cdot P \cdot R}{\beta^2 \cdot P} = R$$

As the weight of *R* goes up, the F-measure begins to approximate simply *R* !

If Precision Very Very Important

Then, as $\beta \to 0$,

$$F_\beta \approx 1 \cdot \frac{P \cdot R}{R}$$

$$= P$$

which is what we expect.