



Text Classification and Naïve Bayes

Precision, Recall, and the F measure



The 2-by-2 contingency table

| | correct | not correct |
|--------------|---------|-------------|
| selected | tp | fp |
| not selected | fn | tn |



Precision and recall

- **Precision:** % of selected items that are correct
Recall: % of correct items that are selected

| | correct | not correct |
|--------------|---------|-------------|
| selected | tp | fp |
| not selected | fn | tn |



A combined measure: F

- A combined measure that assesses the P/R tradeoff is F measure (weighted harmonic mean):

$$F = \frac{1}{\alpha \frac{1}{P} + (1 - \alpha) \frac{1}{R}} = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}$$

- The harmonic mean is a very conservative average; see *IR* § 8.3
- People usually use balanced F1 measure
 - i.e., with $\beta = 1$ (that is, $\alpha = \frac{1}{2}$):

$$F = 2PR/(P+R)$$



Text Classification and Naïve Bayes

Precision, Recall, and the F measure