

Per-Category Evaluation

	c₁	c₂	c₃	...	c_k
d₁	y(+)	y(-)	n(+)		n(+)
d₂	y(-)	n(+)	y(+)		n(+)
d₃	n(+)	n(+)	y(+)		n(-)

How good are the decisions on c_i?



When the system says "yes,"
how many are correct?



Precision = $\frac{TP}{TP + FP}$



Recall = $\frac{TP}{TP + FN}$

Has the category been assigned to
all the docs of this category?

	System ("y")	System ("n")
Human (+)	True Positives TP	False Negatives FN
Human (-)	False Positives FP	True Negatives TN