Evaluating predict

Have test set $\{x_i, y_i\}$, want to evaluate predict method $\hat{y}_i = \begin{cases} 1, & s_i > \tau \\ 0, & s_i \le \tau \end{cases}$

Confusion Matrix

	Predict 1	Predict 0
Actual 1	True Positive (TP)	False Negative (FN)
Actual 0	False Positive (FP)	True Negative (TN)

Predict

1. Calculate score s_i for each i.

2. Order scores : $S_{a_1} \ge S_{a_2} \ge S_{a_3} \ge \dots \ge S_{a_n}$
where $a_1, \dots, a_n$ is permutation of $\{1, \dots, n\}$.

3. Predict $\hat{y}_{a_1}, \dots, \hat{y}_{a_T} = 1$ and $\hat{y}_{a_{T+1}}, \dots, \hat{y}_{a_n} = 0$.

S :	3.1	3	2.8	2.6	2.3	2.1	1.1	1	.8	.6	.5	.2
$\hat{y}$	1	1	1	1	1	1	0	0	0	0	0	0	$\tau = 6$
y	1	0	1	1	0	1	0	1	0	0	1	0

Conf.   TP = 4 | FN = 2
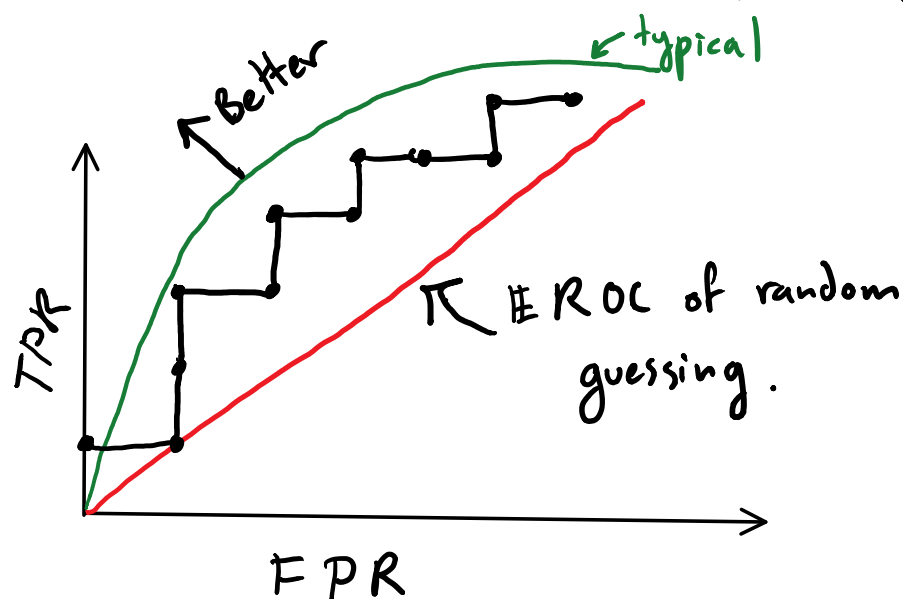
Conf.
Matrix

TP = 4	FN = 2
FP = 2	TN = 4

Reciever - Operating Characteristic (ROC)

True positive rate (TPR) = $\dfrac{TP}{TP + FN}$ } Actual Positive
(aka recall)

False positive rate (FPR) = $\dfrac{FP}{FP + TN}$ } Actual Neg.

T	1	2	3	4	5	6	7	8
TPR	1/6	1/6	2/6	3/6	3/6	4/6	4/6	5/6
FPR	0/6	1/6	1/6	1/6	2/6	2/6	3/6	3/6

Better ← typical

≠ ROC of random guessing.

TPR

FPR

▹ Every element of confusion matrix is represented in ROC

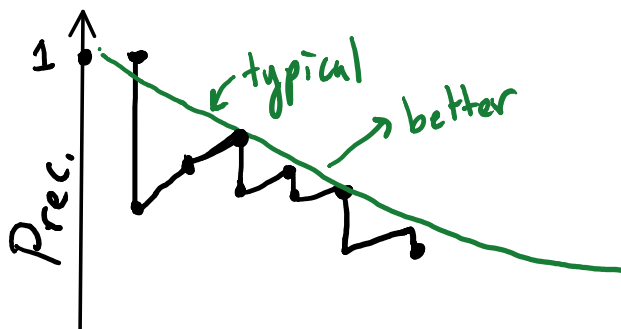## Precision - Recall Curve

In recommendation systems, many 0's — few 1's.

ex ▹ Link prediction in sparse graphs (friend recomm.)
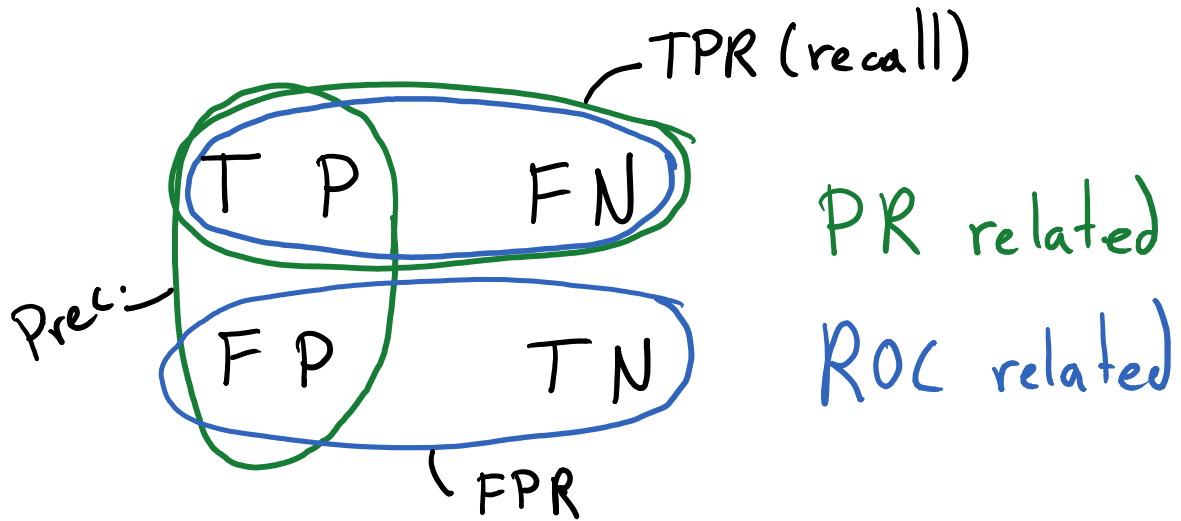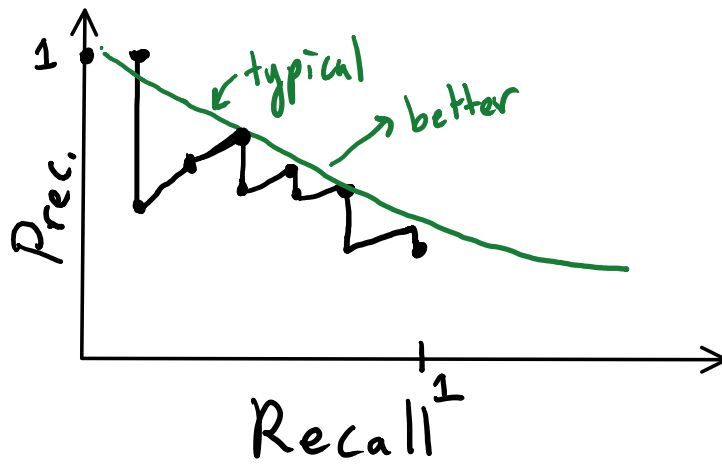 ▹ Search engine (most pages are not relevant)
 ▹ Document retrieval

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

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

$\Big\}$ Neither is sensative to increasing TN

T	1	2	3	4	5	6	7	8	
Recall	$1/6$	$1/6$	$2/6$	$3/6$	$3/6$	$4/6$	$4/6$	$5/6$	...
Prec.	1	$1/2$	$2/3$	$3/4$	$3/5$	$4/6$	$4/7$	$5/8$	...

1

↑ typical → better

Prec.

Recall 1

TPR (recall)

T P   F N

Prec.

F P   T N

FPR

PR related

ROC related

Comparing PR Curves

1

← better at lower recall

Prec.

better at higher recall

Recall 1