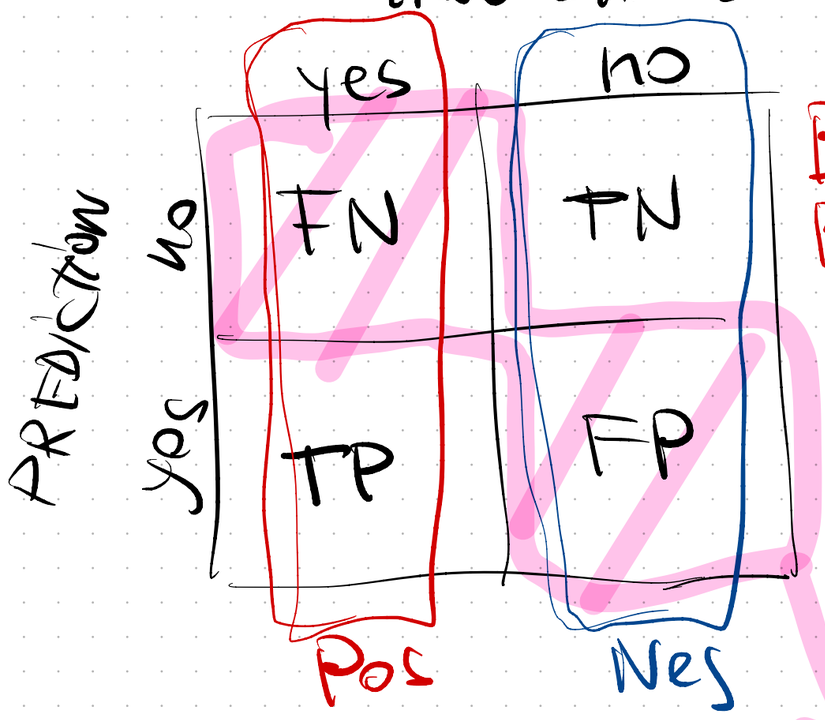
Binary Accuracy ⟹ Confusion Matrix
(2 classes yes/no)

Measure performance ⟹ 4 types of datapoints (counts)



Pos = True Pos : pred = yes, label = yes
False Neg : pred = no, label = yes

Neg = True Neg : pred = no, label = no
False Pos : pred = yes, label = no

mistakes

Often we have score $\Longrightarrow$ prediction $\begin{cases} Y \\ N \end{cases}$? use threshold θ

score $\geq \theta \Rightarrow$ predict yes
score $< \theta \Rightarrow$ predict no

Rank datapoints in test by score.

TEST SET (Z)

r_1 $h(z_1)$ = highest score
r_2 $h(z_2)$ = second highest
$h(z_3)$ = 3rd highest
$h(z_k)$
$h(z_t)$
$h(z_{t+1})$
θ
r_T $h(z_T)$ = lowest score

predict positive

predict negative

$Y(z_t) = yes = pos \Rightarrow TP$
$Y(z_t) = no = neg \Rightarrow FP$
$Y(z_t) = neg \Rightarrow TN$
$Y(z_t) = pos \Rightarrow FN$

$$ACC = \frac{TP + TN}{|\text{test set}|}$$

ACC can be very sensitive to threshold θ !! ⟹ Fix

want: measure of score performance without θ

: score to rank positives higher than negatives

: score(positives) > score(negatives)

ANSWER: ROC curve = curve for all thresholds θ

$h(z_1)$ highest score

$h(z_4)$ = lowest score

θ

TP_θ

FP_θ

TN_θ

FN_θ

|test set| = $Z = TP_\theta + FP_\theta + TN_\theta + FN_\theta$ (any θ)

For thresh θ

TPR = true pos rate = recall

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

FPR = false pos rate = type I err

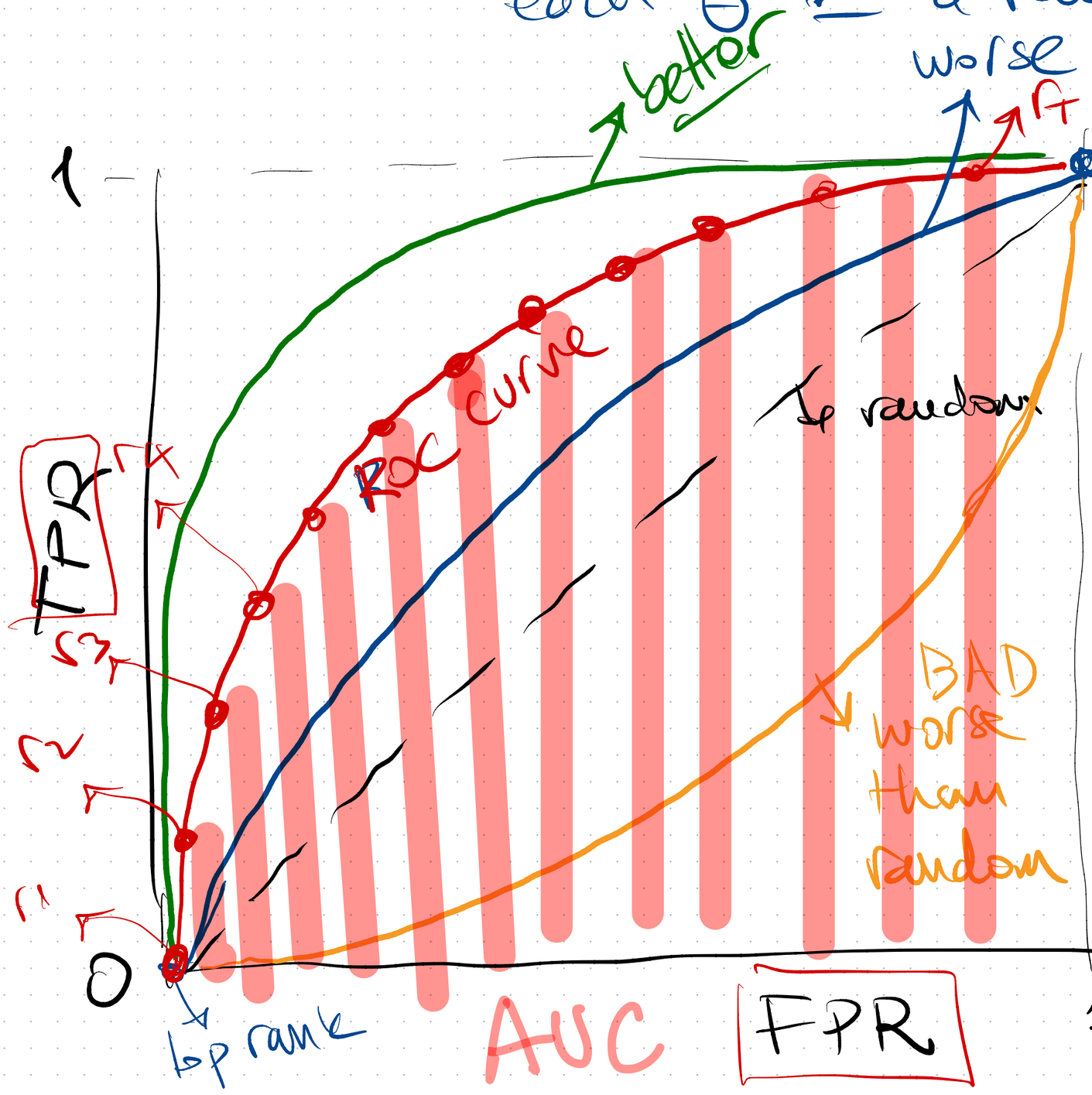
"fall out"

$= \frac{FP}{neg} = \frac{FP}{FP+TN}$

Plot TPR (y axis) vs FPR (x axis) for all thresholds θ

Hidden on plot

each θ ≈ a rank $\{1, 2, \ldots, T\}$



better: TPR rises fast vs FPR

worse: TPR rises slow vs FPR

random predict TPR rises lin FPR

BAD: worse than random

Good performance ROC = green (rise fast)

MEASURE: Area Under Curve AUC

max = 1

RANKING by score

$h(z_1)$

$h(z_2)$ → negative $Y(z_2) = N$

⋮

$h(z_7)$ → positive $Y(z_7) = Y$

⋮

$h(z_T)$

} inversion (z_2, z_7) rank incorrectly.

$$\#\text{pairs total} = \frac{T(T-1)}{2} = \#\text{correct} + \#\text{inversions}$$

ex: AUC ~ $\frac{\#\text{correct ranked pairs}}{\#\text{total pairs}}$

HW2A pb 5: AUC $\approx$ $\frac{\#\text{correct pairs (in correct order)}}{\#\text{total pairs}}$ = % Correct Ranked pairs

Score-sort Ranking (highest = top) labels = $y_1 y_2 \dots y_N$ binary $\in \{+, -\}$

Ranks score Label

K=1 $y_1 = \oplus$ → TP

K=2 $y_2 = \oplus$ → TP

K=3 $y_3 = \ominus$ → FP

$y_4 = \oplus$ → TP

K — — — y_K

K+1 — — — y_{K+1}

N-1 $y_{N-1} = \oplus$ → false neg FN

N — — — $y_N = \ominus$ → true neg TN

θ = threshold, pred $+$ / $-$ between ranks K, K+1

At threshold θ for scores $\Leftrightarrow$ k for ranks

$$TPR_k = \frac{TP_{(k)}}{\text{all pos}} = \frac{\sum_{i=k}^{top} [y_i = +]}{N_+}$$ 1 or 0

$$FPR_k = \frac{FP_k}{\text{total neg}} = \frac{\sum_{i=k}^{top} [y_i = -]}{N_-}$$

N_+ → total pos

1 if $y = -$, 0 if $y = +$

#labels negative above threshold

$N_+ = TP + FN$ = total pos

$N_- = TN + FP$ = total neg

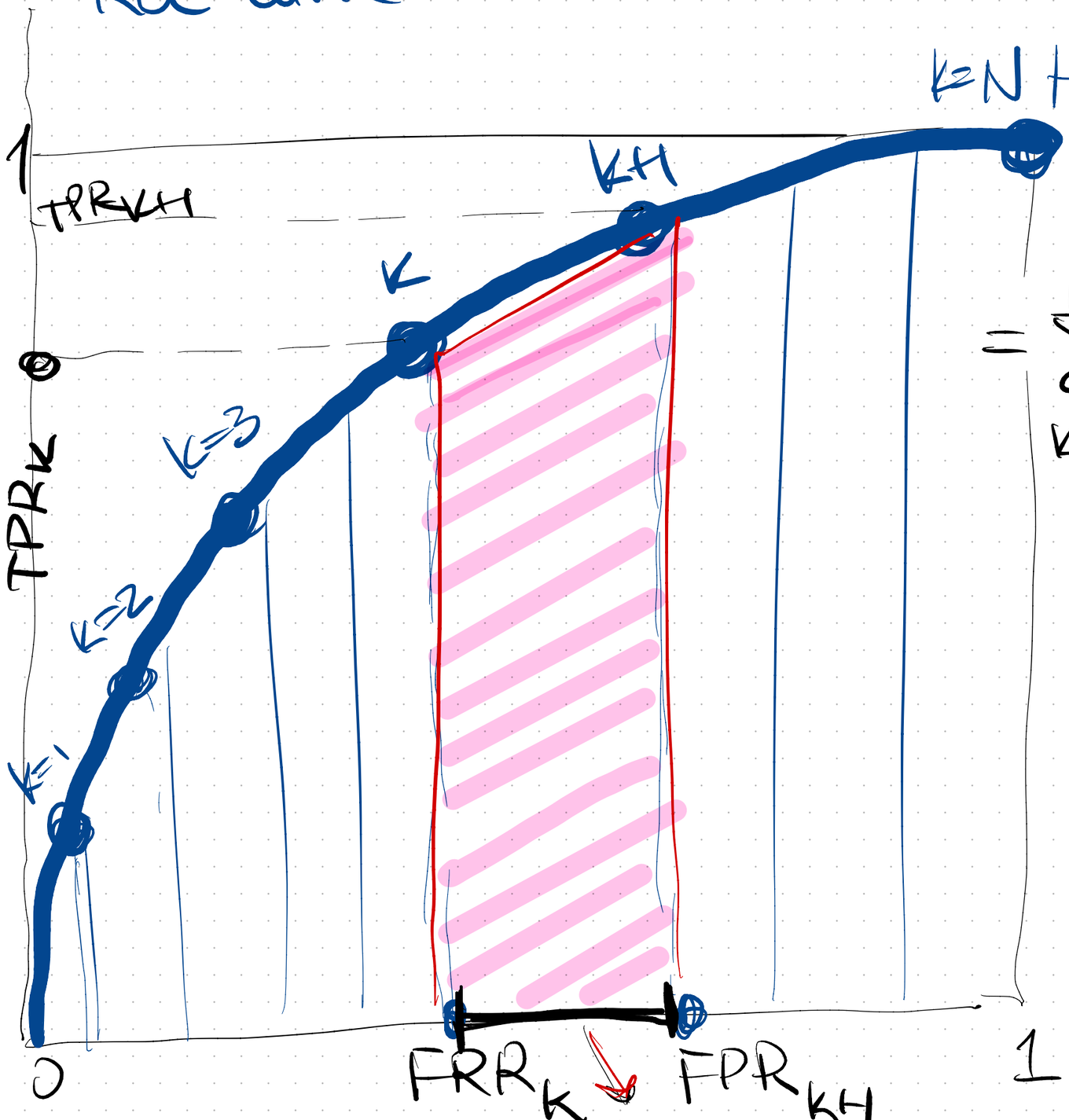
$N_+ + N_- = N$

threshold at rank N $\Leftarrow$ above = all (nothing below)

TP

FP

ROC curve



$$\text{AUC} \simeq \sum \text{area-trapezoids}$$

$$= \sum_{k=1}^{N} \left(\frac{TPR_k + TPR_{k+1}}{2} \right) \cdot \left(FPR_{k-1} - FPR_k \right)$$

height · base

$$= \sum_{k=1}^{N} \frac{\sum_{i=k}^{N} [y_i = +] + \frac{1}{2}[y_{k+1} = +]}{N_+} \cdot \frac{[y_{k+1} = -]}{N_-}$$

- - - - -
- - - - -
- - - - -

want "prop of correct rank pairs"

$$\frac{1}{N_+ N_-} \underbrace{\sum_{k=1}^{N} \sum_{j:k=-}}_{\text{every neg}} \underbrace{\sum_{i=k+1}^{\text{top}}}_{\text{ranked above k}} [y_i = +]$$

all pairs (+/−)

pos