

OUT-OF-DISTRIBUTION

AND

ANOMALY DETECTION

IA Frameworks

Joseba Dalman

LEARNING OUTCOMES

- Master the fundamental OOD / AD concepts
- Understand the main algorithm families, their advantages and weaknesses
- Familiarize with the most common evaluation metrics

WHAT IS AN ANOMALY ?

Anomaly

"An observation that deviates considerably from some concept of normality"

WHY CARE ABOUT ANOMALY DETECTION?

- Intrusion detection (cybersecurity)
- Fraud detection (finance)
- Industrial fault / damage detection
- Medical diagnosis
- [NEW!] Out-of-Distribution
detection for ML and DL models

MAIN CHALLENGES

- Ill-defined problem
- Absence of abnormal data
- Deep Learning use-cases:
data is very high-dimensional

WHAT IS AN ANOMALY ?

"An observation that deviates considerably from some concept of normality"

A typical mathematical formalization :

$X \subset \mathbb{R}^D$ data space

\mathbb{P}^+ ground-truth law of normal data on X with p.d.f. p^+

$A_z = \{x \in X \mid p^+(x) \leq z\} \leftarrow$ set of anomalies
↑?

CAREFUL!

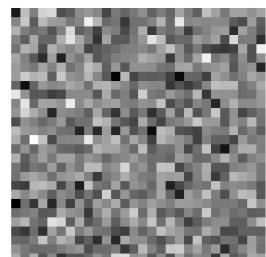
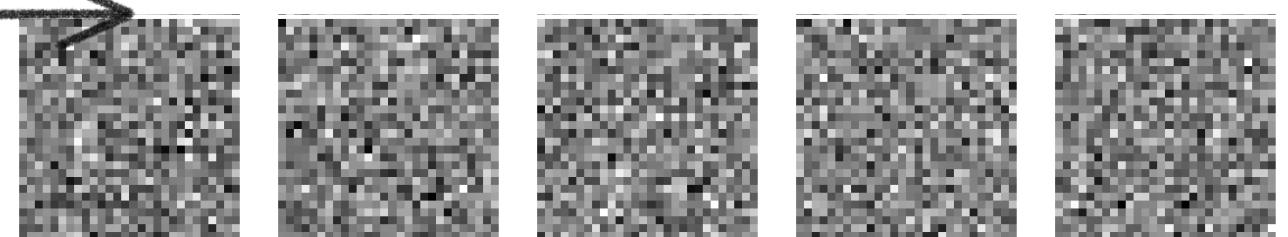
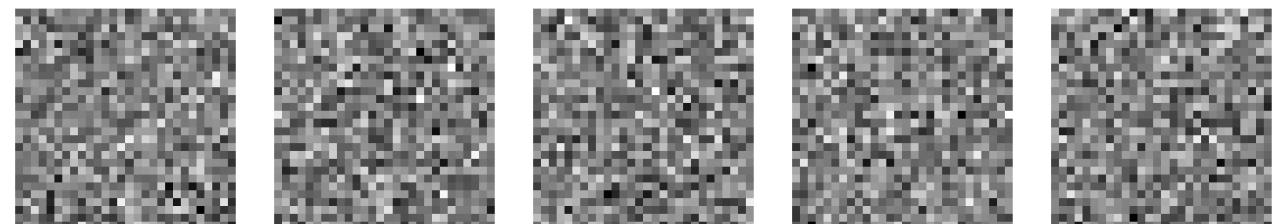
What is the concept of normality
used in this formalization ?

CAREFUL!

What is the concept of normality used in this formalization ?

10 NORMAL

images



Question : which of these
2 images is normal ?

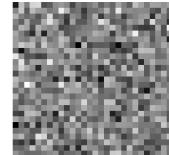
CAREFUL!

Yet: pixels are i.i.d.

$N(0,1)$ random variables :

\Rightarrow Likelihood Ratio

$$\frac{f(\text{[]})}{f(\text{[]})} = \prod_{ij} \frac{1}{e^{-\frac{x_{ij}^2}{2}}}$$

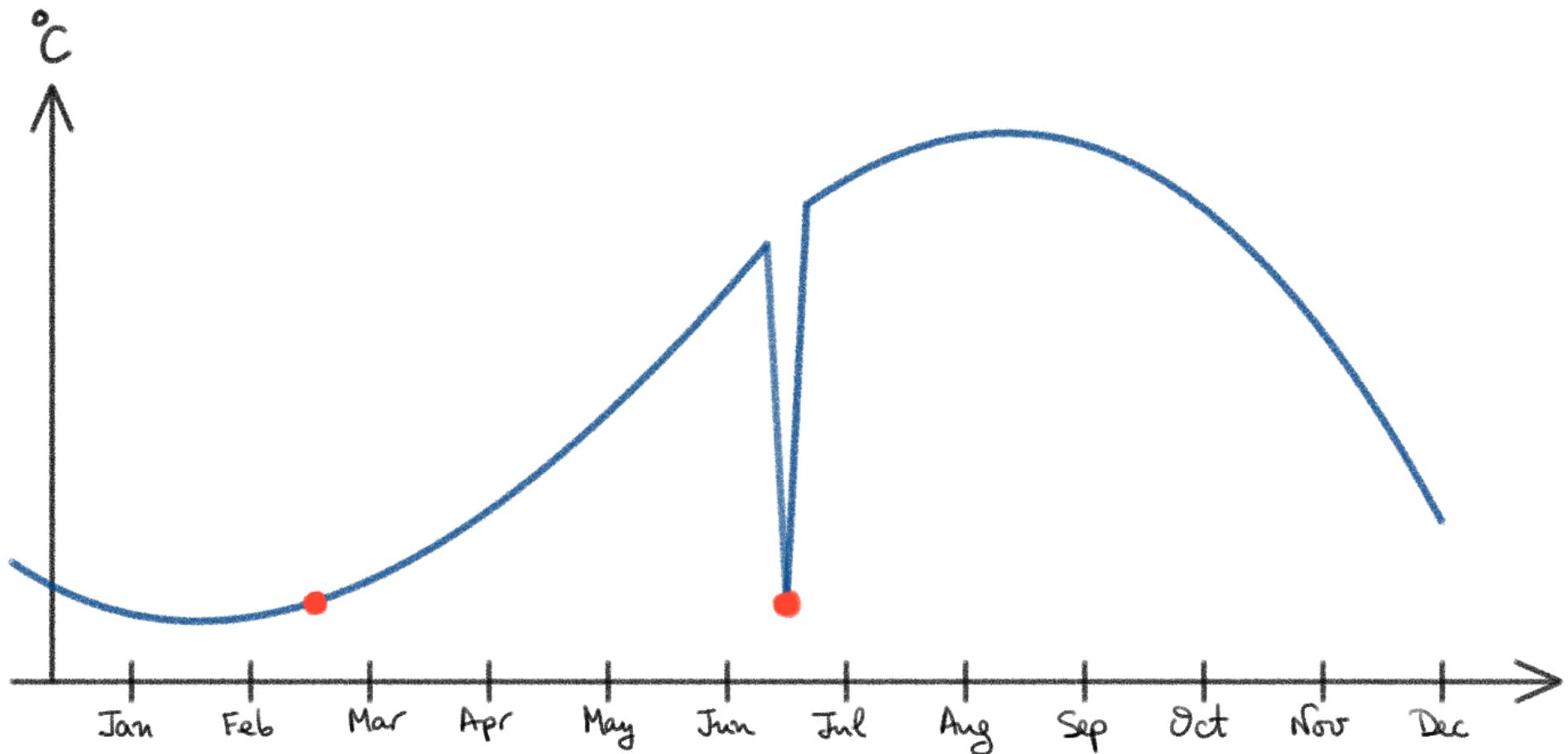
value of pixel ij
in 

TYPES OF ANOMALIES

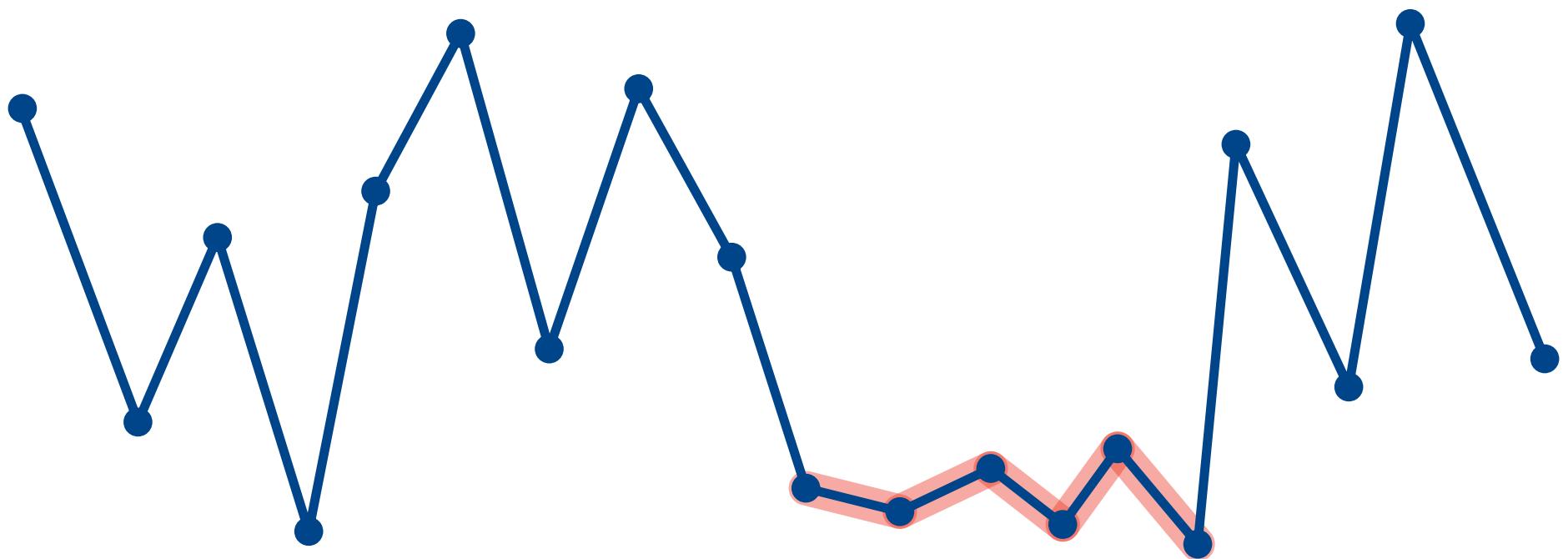
Traditionally:

- Point anomalies
- Conditional / Contextual anomalies
- Group / Collective anomalies

CONTEXTUAL ANOMALIES



COLLECTIVE ANOMALIES

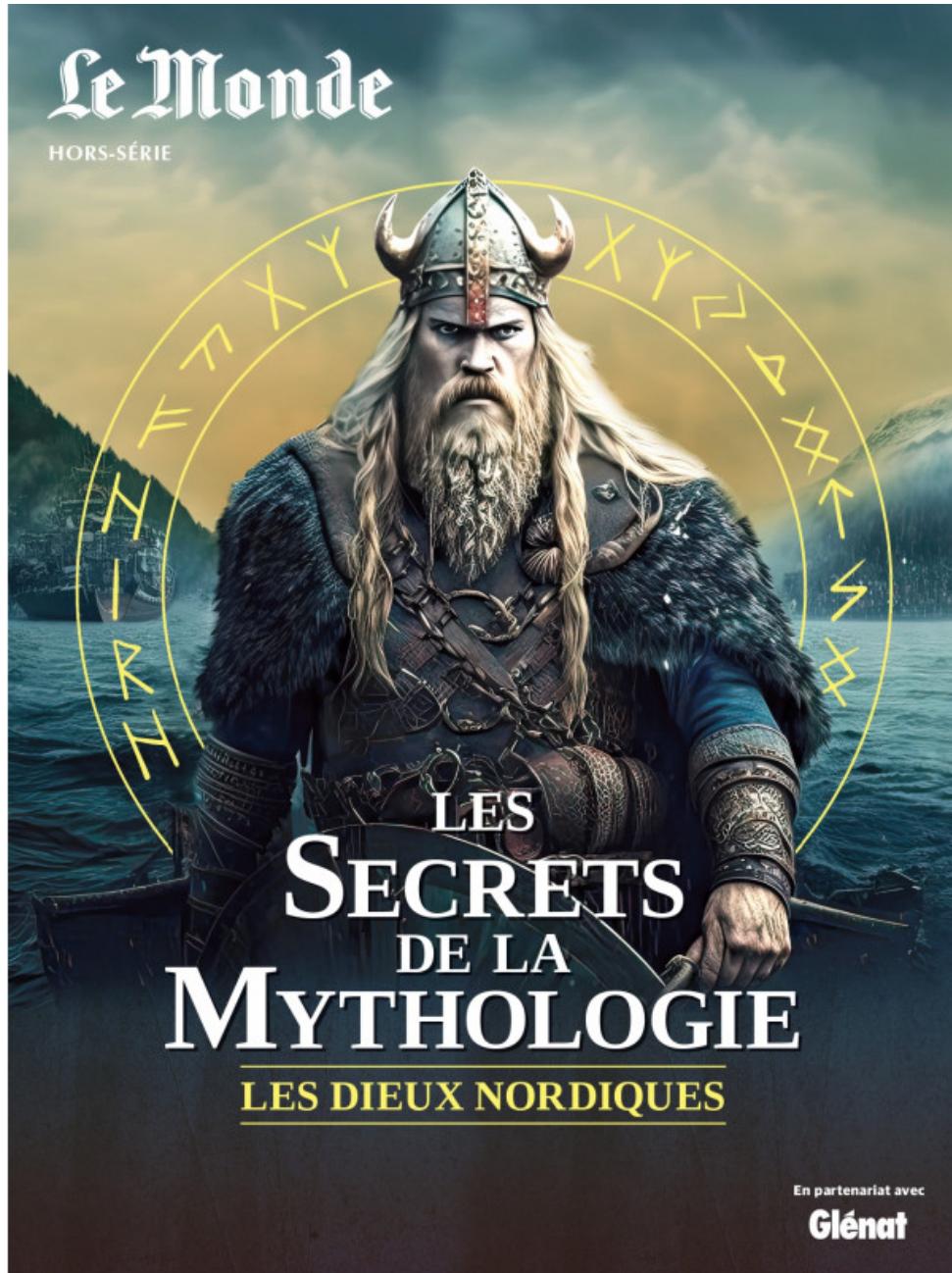


TYPES OF ANOMALIES

Since Deep Learning :

- Low-level sensory anomalies
- High-level semantic anomalies

LOW LEVEL, SENSORY



The book is interesting and it conveys a meaningful message.

HIGH LEVEL, SEMANTIC



The girl was
playing with a
colorless green
liquid.

TYPES OF ANOMALIES

Anomaly not drawn from \mathbb{P}^+

Outlier drawn from \mathbb{P}^+ w low proba

Novelty from an evolving \mathbb{P}^+

ANOMALY DETECTION STRATEGIES

→ Model-agnostic

Given $X_1, \dots, X_n \sim \mathbb{P}^+$, $X_{\text{test}} \in A_c ?$

→ Model-based

Given $(X_1, Y_1), \dots, (X_n, Y_n) \sim \mathbb{P}^+ \times \mathbb{P}_Y$

and $h: X \rightarrow Y$, $X_{\text{test}} \in A_c ?$

ANOMALY DETECTION ALGORITHMS

AD algorithm

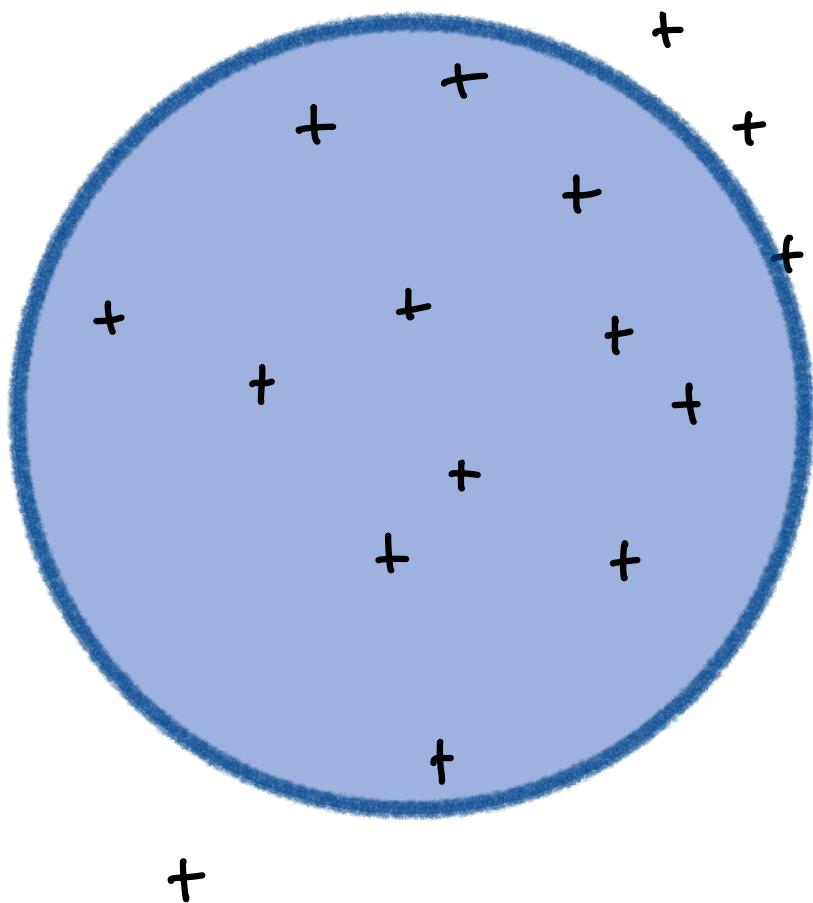
$s: X \rightarrow \mathbb{R}$ and $\tau \in \mathbb{R}$
such that

- { if $s(x) \leq \tau \Rightarrow x$ is declared "normal"
- if $s(x) > \tau \Rightarrow x$ is declared "anomalous"

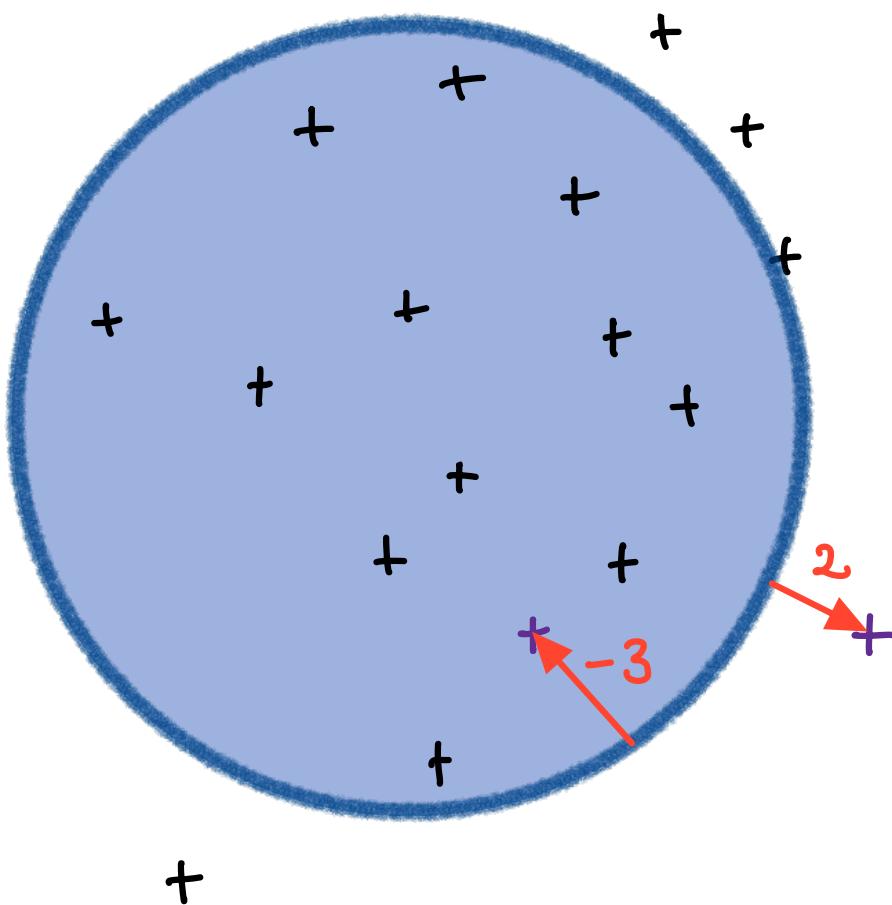
ANOMALY DETECTION ALGORITHMS

- One-class classifier
- Probability density estimation
- Reconstruction
- Distance
 - + predictor $h: X \rightarrow Y$
 - (if available)

ONE-CLASS SVM



ONE-CLASS SVM



$s(x)$ = signed distance
from x to

$$\bar{z} = 0$$

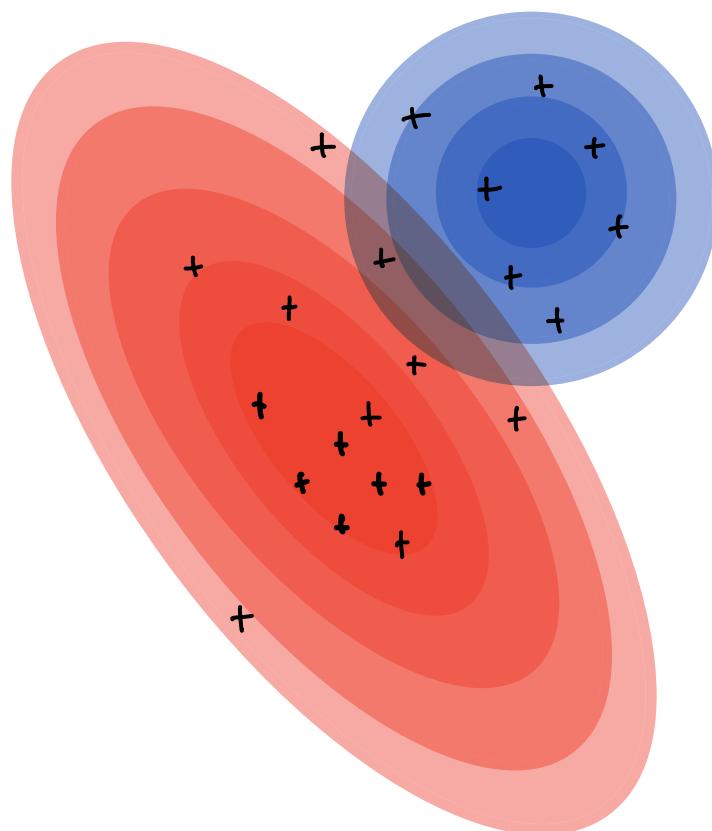
$\Rightarrow \begin{cases} x \in \textcircled{O} \Rightarrow \text{"normal"} \\ x \notin \textcircled{O} \Rightarrow \text{"anomaly"} \end{cases}$

PROBABILITY DENSITY ESTIMATION

e.g. with a Gaussian Mixture Model (GMM)

i.e. $P^+ \simeq \sum_{i=1}^k w_i N(\mu_i, \Sigma_i)$

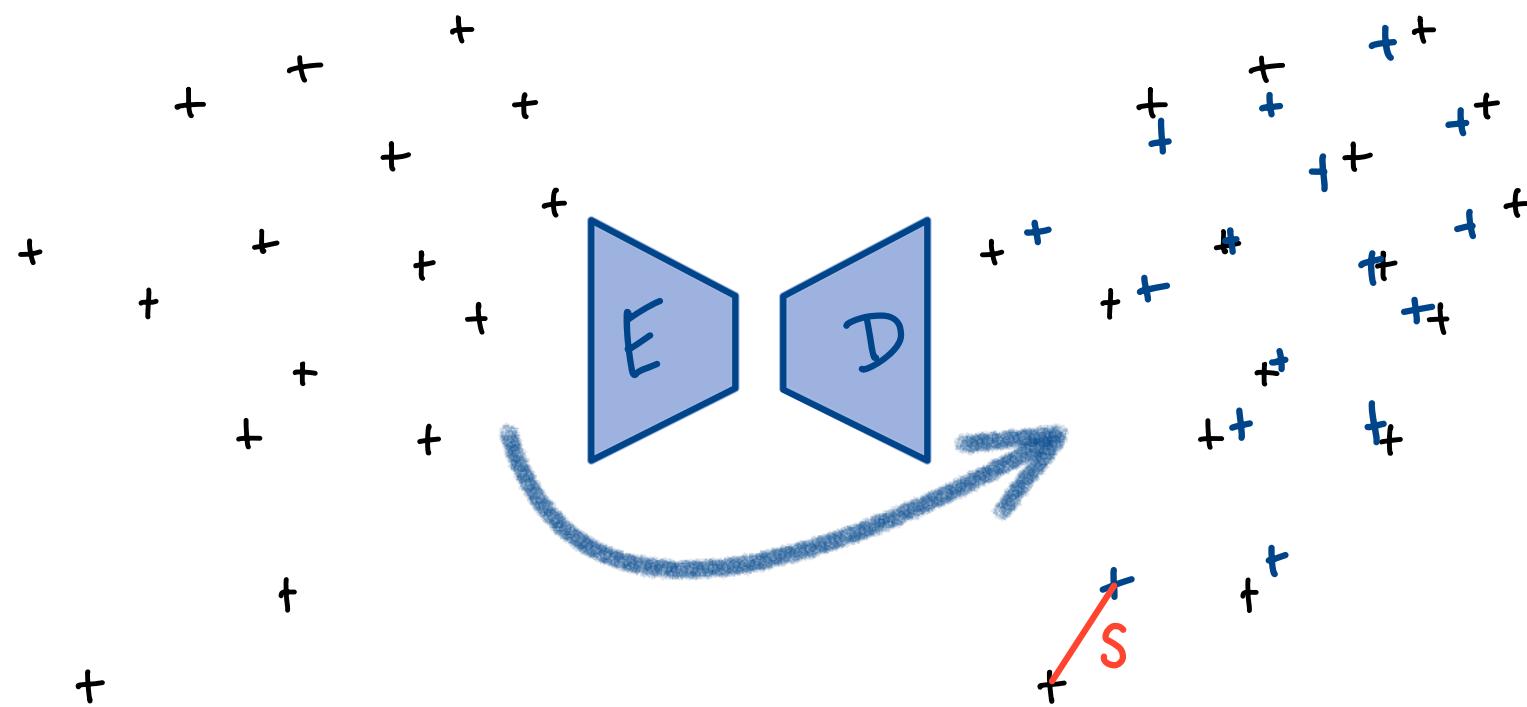
GMM fitted on
 x_1, \dots, x_n



$$S(x) = -\text{log-likelihood}$$

of x under the GMM

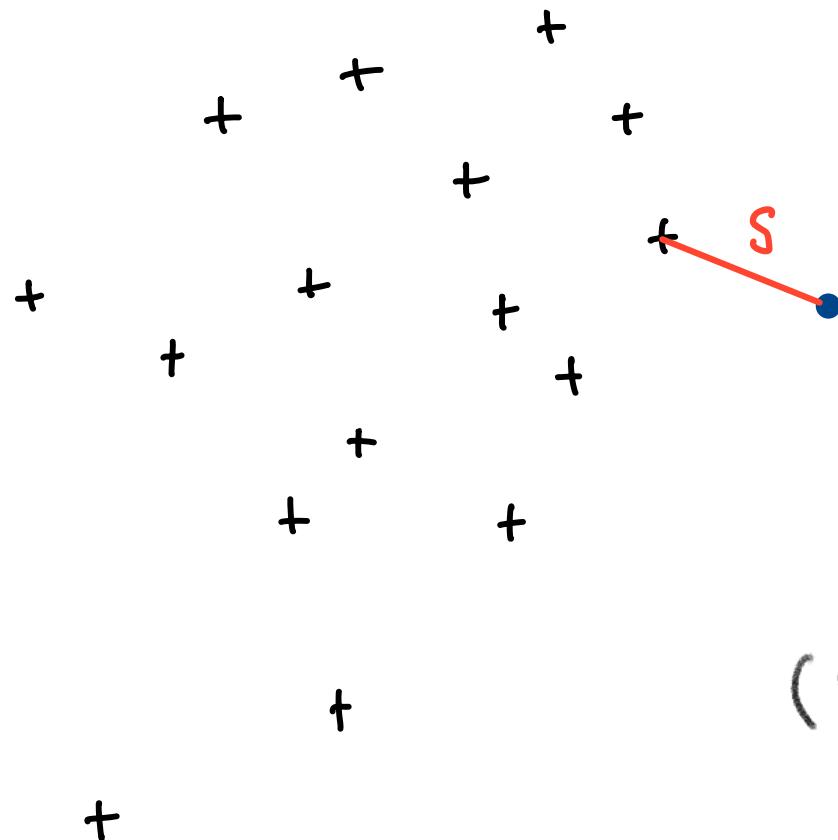
RECONSTRUCTION w. AUTOENCODERS



$$s(x) = d(x, D(E(x)))$$

DISTANCE w. K-NEAREST NEIGHBORS

$+ \sim \mathbb{P}^+$



$s(x) = \text{distance to}$
 $(k\text{-th}) \text{ nearest neighbor}$

QUIZ

	sensitive to outliers ?	scales to large dim?	≈ easy to pick ?
SVM			
GMM			
AE			
K-NN			

EVALUATION METRICS

Normal = -

Anomaly = +

		Predicted	
		-	+
Ground Truth	-	True Negative	False Positive
	+	False Negative	True Positive

Confusion Matrix

EVALUATION METRICS

For a fixed threshold z :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

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

		Predicted	
		-	+
Ground Truth	-	True Negative	False Positive
	+	False Negative	True Positive

Recall /

$$\text{True Positive Rate} = \frac{TP}{TP + FN}$$

EVALUATION METRICS

For a fixed threshold z :

$$F_1 \text{ score} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}$$

$$\begin{aligned} \text{False Positive} \\ \text{Rate} \end{aligned} = \frac{FP}{TN + FP}$$

Quiz

1. Is accuracy an adequate metric for AD? Why?
2. Give an example of a confusion matrix where precision and recall are both high.
3. Express the F_1 score in terms of the elements of the confusion matrix
4. The F_β score is $F_\beta = (1+\beta^2) \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \beta^2 \text{Recall}}$
What term do we give more importance to
(i) if $\beta < 1$? (ii) if $\beta > 1$?

EVALUATION METRICS

Threshold independent: AUC-ROC / AUROC

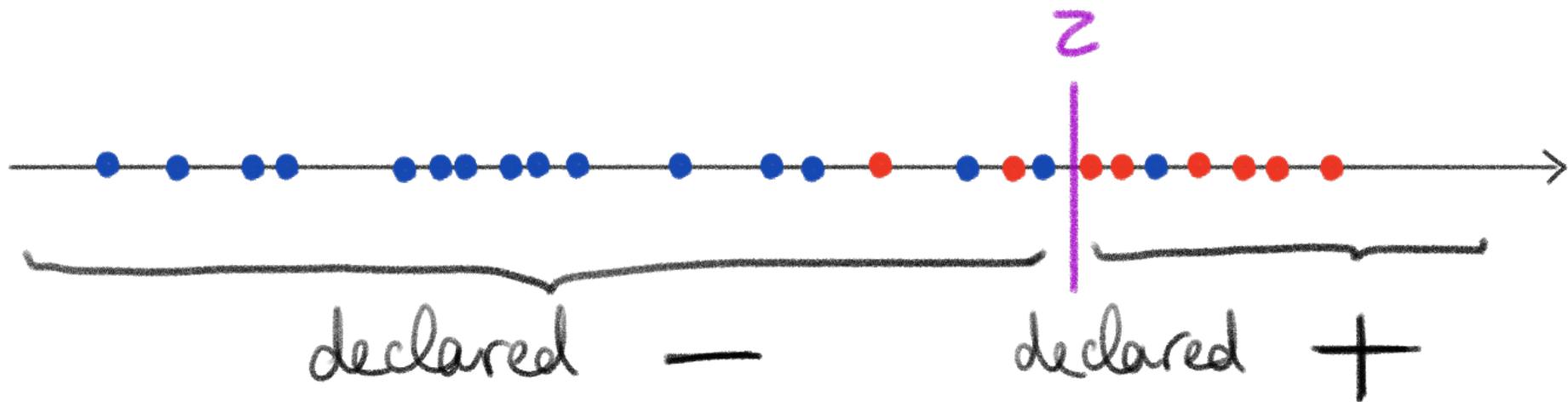
ROC = "Receiver Operating Characteristic"

Given: $s: X \longrightarrow \mathbb{R}$

- = normal (ground-truth)
- = anomaly (ground-truth)



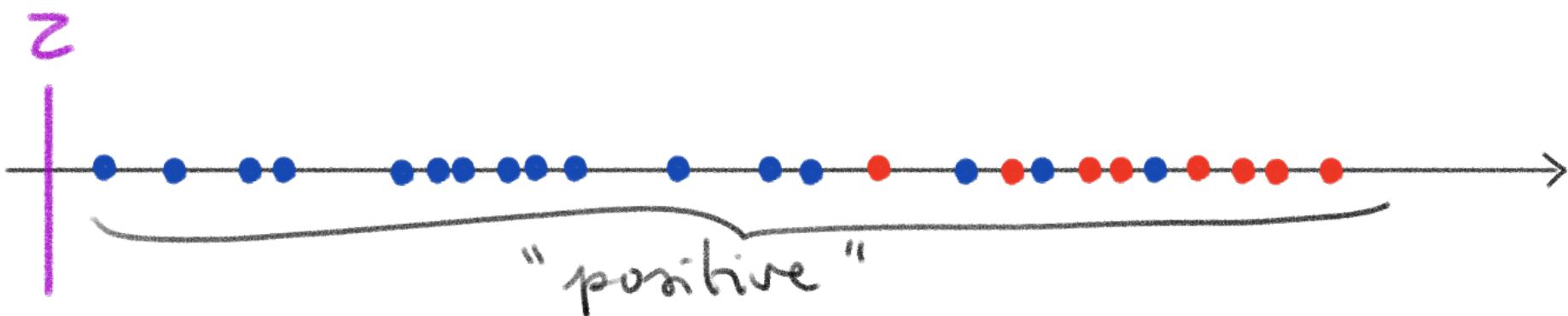
EVALUATION METRICS



$$TPR(\tau) = \frac{\text{* of } \textcolor{red}{\bullet} \text{ and } +}{\text{* of } \textcolor{red}{\bullet}}$$

$$FPR(\tau) = \frac{\text{* of } \textcolor{blue}{\bullet} \text{ and } -}{\text{* of } \textcolor{blue}{\bullet}}$$

EVALUATION METRICS

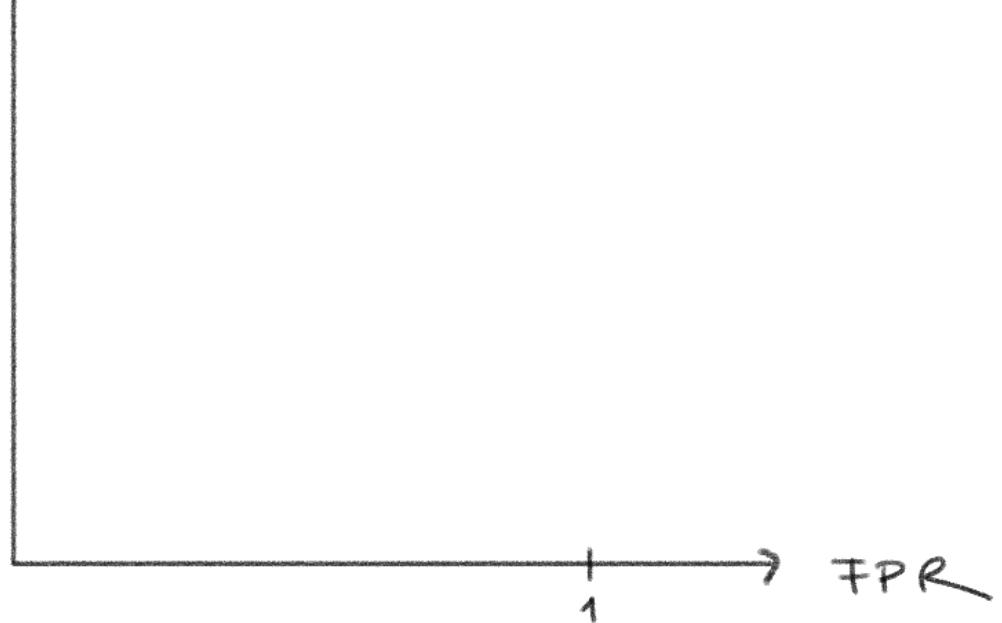


TPR

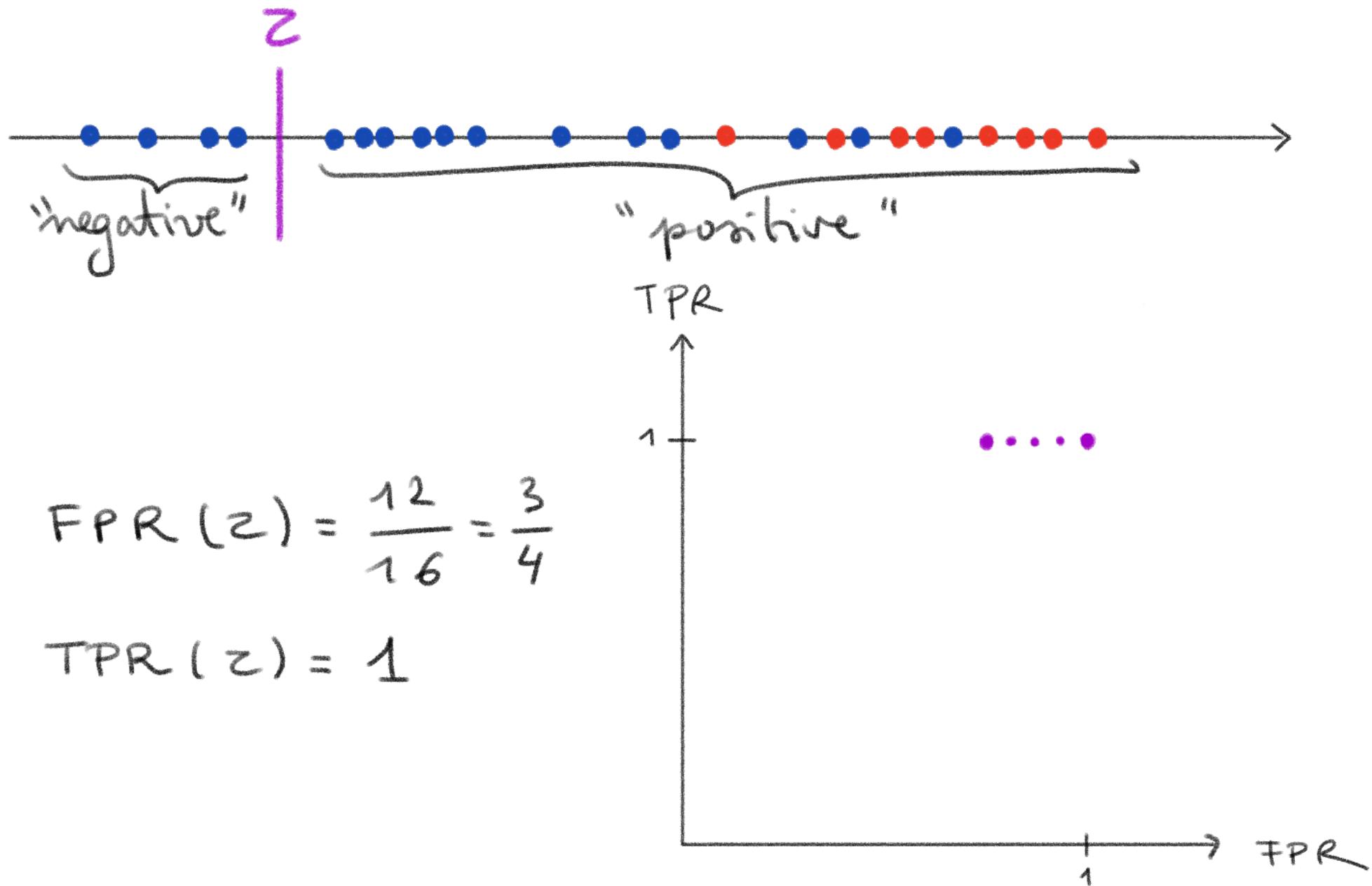


$$FPR(z) = 1$$

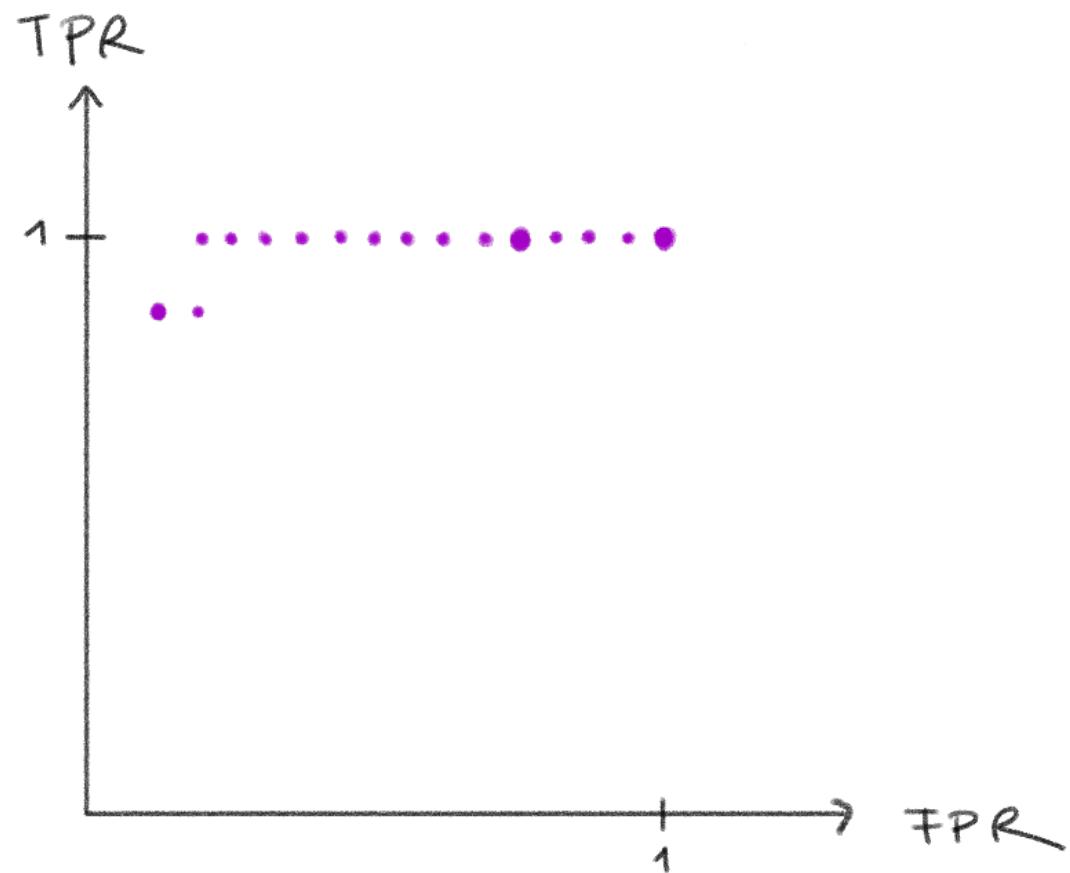
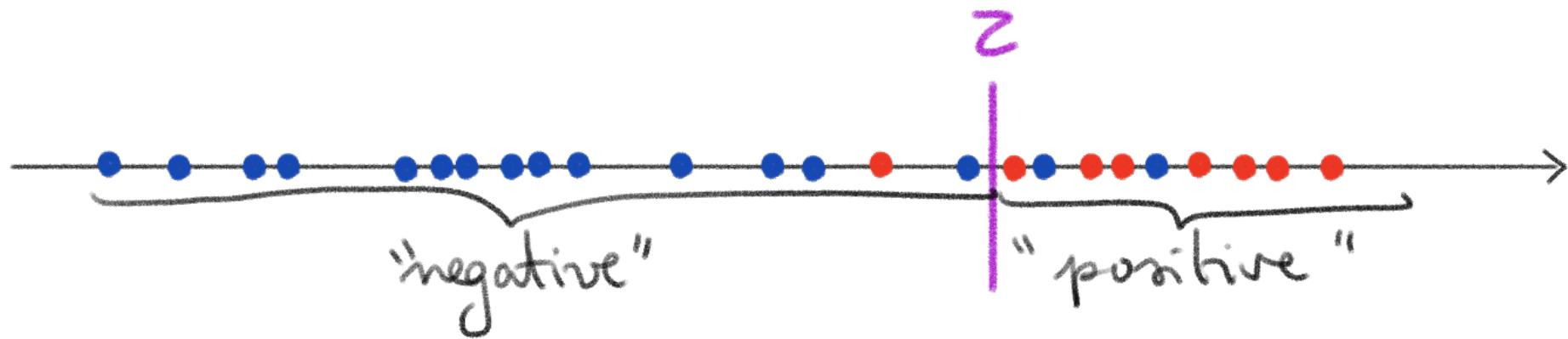
$$TPR(z) = 1$$



EVALUATION METRICS



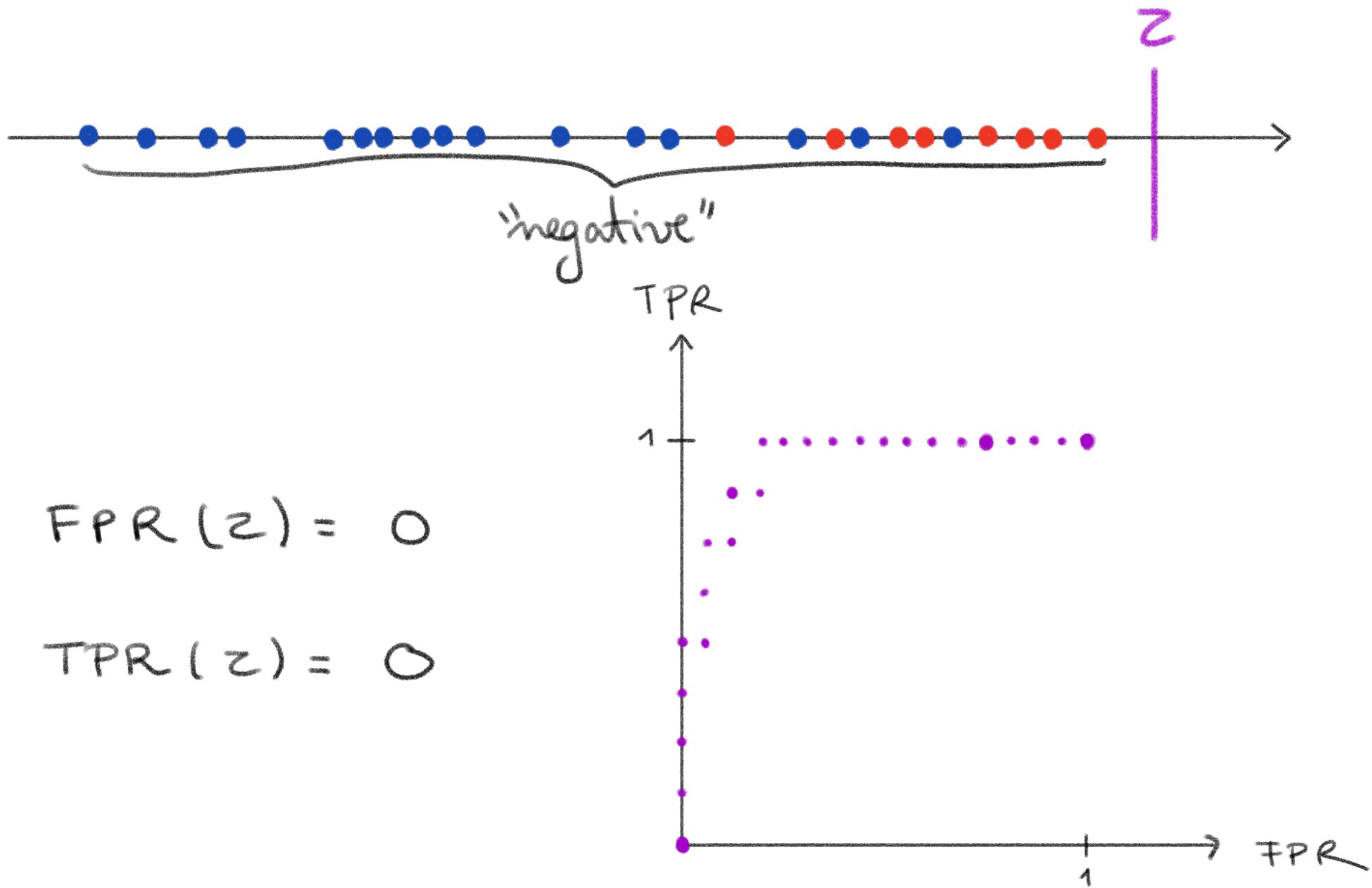
EVALUATION METRICS



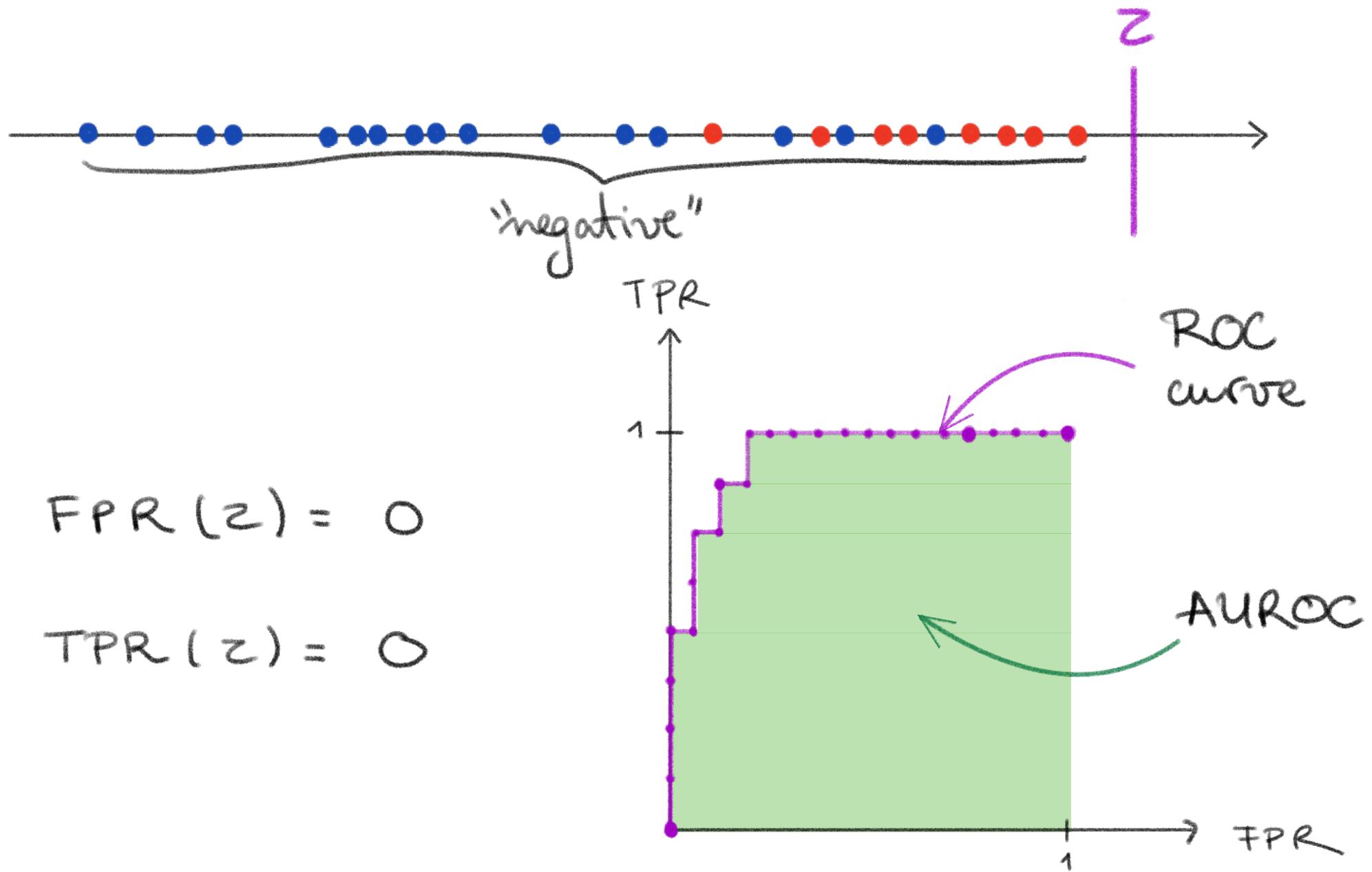
$$FPR(z) = \frac{1}{8}$$

$$TPR(z) = \frac{7}{8}$$

EVALUATION METRICS



EVALUATION METRICS



QUIZ

1. What would a "perfect" score look like?
And the ROC curve/AUROC associated?
2. Same question for a "random" score.
3. Same question for a "bad" score.
4. Can we generalize the ROC curve /AUROC to the case of probability distributions?

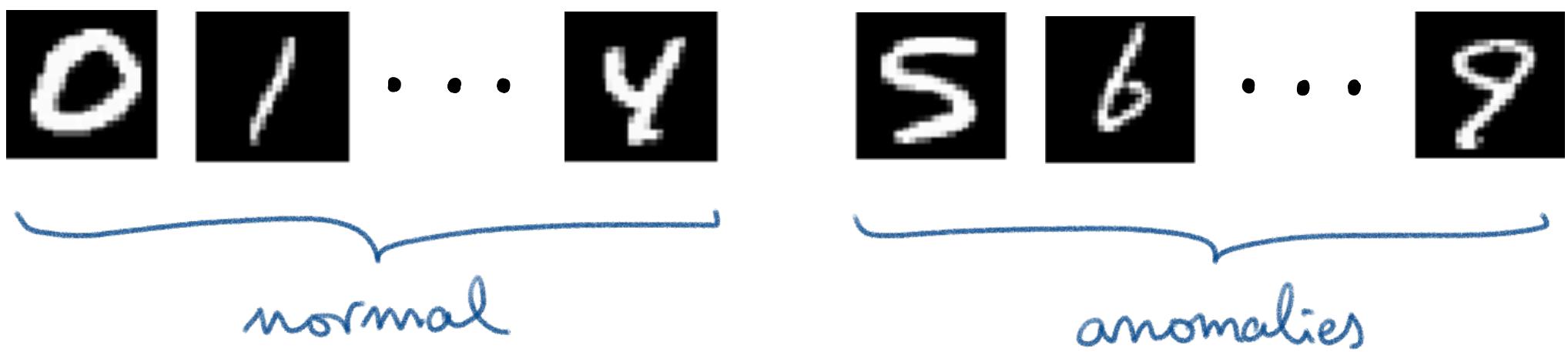
BENCHMARKING



Lack of datasets for high-dimensional
data (images, text)

see for example MVTecAD

Solution 1 .



BENCHMARKING



Lack of datasets for high-dimensional
data (images, text)

see for example MVTecAD

Solution 2 .

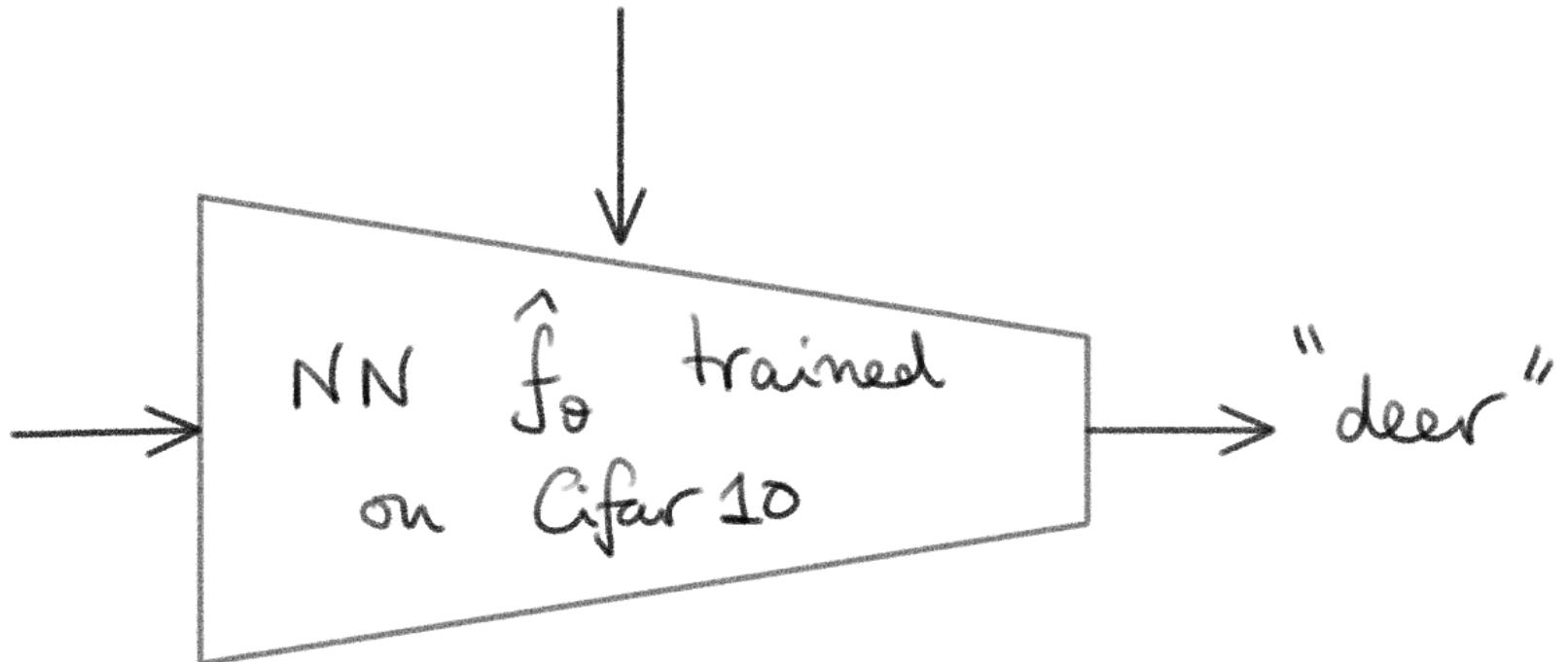


MNIST → normal



CIFAR10 → anomaly

OUT-OF-DISTRIBUTION DETECTION



OoD = AD ?

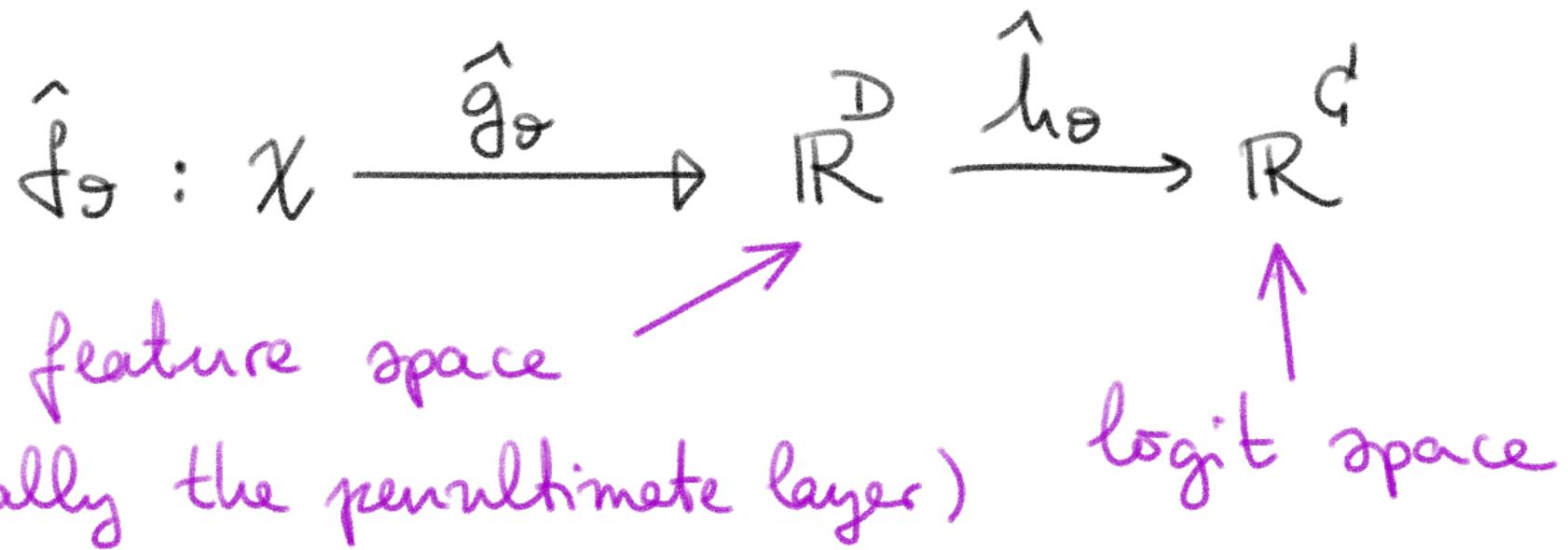
→ They accept similar mathematical formulations, but:

1. In OoD detection the space of anomalous images can be much larger

2. In OoD detection we have access to feature representations via \hat{f}_θ

OoD Scores

for a classifier \hat{f}_0 :



→ logit-based scores: $s(x) = \hat{s}(\hat{f}_\theta(x))$

→ feature-based scores: $s(x) = \tilde{s}(\hat{g}_\theta(x))$

LOGIT-BASED SCORES

→ Maximum Logit Score MLS

$$\tilde{s}(z) = - \max_{1 \leq c \leq C} z_c$$

→ Maximum Softmax Probability

$$\tilde{s}(z) = - \max_{1 \leq c \leq C} P_c(z)$$

LOGIT-BASED SCORES

→ Energy

$$\tilde{S}(z) = -T \log \sum_{c=1}^C e^{z_c/T}$$

tunable
hyperparameter

→ Entropy

$$\tilde{S}(z) = - \sum_{c=1}^C p_c(z) \log p_c(z)$$

FEATURE-BASED SCORES

→ Deep K-nearest neighbor (DKNN)

Given z_1, \dots, z_N and z^* :

\uparrow \uparrow \uparrow
fitting dataset test point

(i) Compute $d_1 = d(z_1, z^*), \dots, d_N = d(z_N, z^*)$

(ii) Rank $d_{(1)} \leq d_{(2)} \leq \dots \leq d_{(N)}$ hyper-parameter

(iii) DKNN score $\tilde{s}(z^*) = d_{(k)}$

FEATURE-BASED SCORES

→ Mahalanobis

Given z_1, \dots, z_N and z^* :

\uparrow \uparrow \uparrow
fitting dataset test point

(i) Fit G Gaussians: $\mathcal{N}(\mu_c, \Sigma)$, $c=1, \dots, G$

$$\hat{\mu}_c := \frac{1}{N_c} \sum_{i:y_i=c}^N z_i, \quad \hat{\Sigma} = \frac{1}{N} \sum_{c=1}^G \sum_{i:y_i=c} (z_i - \hat{\mu}_c)(z_i - \hat{\mu}_c)^T$$

FEATURE-BASED SCORES

→ Mahalanobis

(ii) Mahalanobis score:

$$\tilde{s}(z) = \max_{1 \leq c \leq C} (z - \hat{\mu}_c)^T \hat{\Sigma}^{-1} (z - \hat{\mu}_c)$$

Exercise : If $C=1$, $\hat{\mu}_c = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\hat{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$,
draw $A \subseteq \mathbb{R}^2$ given by $A = \{z : \tilde{s}(z) \leq 1\}$

RESOURCES

Surveys

1. A unifying review of deep and shallow anomaly detection,
Ruff et al. 2021



2. Generalized out-of-distribution detection, Yang et al. 2021



RESOURCES

Github repositories

1. OpenOOD



2. ADBench



CHALLENGES

- Representation learning for AD
- Characterization of anomalies
- Interpretability of AD
- Creation of a unified methodology
for AD research:
benchmarks, metrics,...

Kahoot
time!

Quiz

1. Why is AD important for ML/DL?
2. Name 2 classes of AD algorithms.
3. Write down the formula for 1 metric.
4. Write down 1 sentence stating 1 thing you understood well.
5. Write down 1 question with something you would like to know more about.