

Machine Learning 5

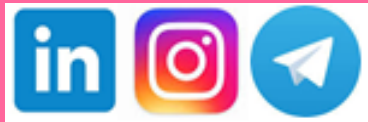
Evaluation Metrics

Accuracy – Sensitivity – Specificity

Precision – F1 Score



behzad.amanpour





Accuracy

$$\frac{\text{True Decisions}}{\text{All Decisions}}$$

- **Imbalanced Data**
 - 90 normal cases (negative class)
 - 10 cancerous cases (positive class)



Sensitivity - Specificity

- **Sensitivity, Recall, True Positive Rate**

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

- **Specificity, Selectivity, True Negative Rate**

$$\text{TNR} = \frac{\text{TN}}{\text{N}} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$



Sensitivity - Specificity

- **Precision, Positive Predictive Value**

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

- **F1 Score (Harmonic Mean of Precision and Sensitivity)**

$$F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$



Example

20 samples: 10 cancer + 10 normal

- **Algorithm1:**

- **all predicted cancer**

$$\text{TPR} = \frac{\text{TP}}{\text{P}} = 100\%$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} = 50\%$$

$$\text{TNR} = \frac{\text{TN}}{\text{N}} = 0\%$$

- **Algorithm2:**

- **5 cancerous samples predicted correctly**

- **15 samples predicted normal**

$$\text{TPR} = \frac{\text{TP}}{\text{P}} = 50\%$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} = 100\%$$

$$\text{TNR} = \frac{\text{TN}}{\text{N}} = 100\%$$



WIKIPEDIA
The Free Encyclopedia

[Main page](#)

[Contents](#)

[Current events](#)

[Random article](#)

[About Wikipedia](#)

[Contact us](#)

[Donate](#)

[Contribute](#)

[Help](#)

[Learn to edit](#)

[Community portal](#)

[Recent changes](#)

[Upload file](#)

[Tools](#)

Wikipedia

Article [Talk](#)

Precision and recall

From Wikipedia, the free encyclopedia

In [pattern recognition](#), [information retrieval](#) and [classif](#) from a [collection](#), [corpus](#) or [sample space](#).

Precision (also called [positive predictive value](#)) is the the fraction of relevant instances that were retrieved. f

Consider a computer program for recognizing dogs (th twelve dogs, the program identifies eight dogs. Of the cats (false positives). Seven dogs were missed (false (true positives / selected elements) while its recall is 5

When a [search engine](#) returns 30 pages, only 20 of w tells us how valid the results are, while its recall is 20/

Adopting a hypothesis-testing approach from [statistics](#) [type I and type II errors](#) (i.e. perfect [specificity](#) and ser (no false negative).

[sensitivity](#), [recall](#), [hit rate](#), or [true positive rate](#) (TPR)

$$\text{TPR} = \frac{\text{TP}}{\text{P}} = \frac{\text{TP}}{\text{TP} + \text{FN}} = 1 - \text{FNR}$$

[specificity](#), [selectivity](#) or [true negative rate](#) (TNR)

$$\text{TNR} = \frac{\text{TN}}{\text{N}} = \frac{\text{TN}}{\text{TN} + \text{FP}} = 1 - \text{FPR}$$

[precision](#) or [positive predictive value](#) (PPV)

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} = 1 - \text{FDR}$$

[negative predictive value](#) (NPV)

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} = 1 - \text{FOR}$$

[miss rate](#) or [false negative rate](#) (FNR)

$$\text{FNR} = \frac{\text{FN}}{\text{P}} = \frac{\text{FN}}{\text{FN} + \text{TP}} = 1 - \text{TPR}$$

[fall-out](#) or [false positive rate](#) (FPR)

$$\text{FPR} = \frac{\text{FP}}{\text{N}} = \frac{\text{FP}}{\text{FP} + \text{TN}} = 1 - \text{TNR}$$

[false discovery rate](#) (FDR)

$$\text{FDR} = \frac{\text{FP}}{\text{FP} + \text{TP}} = 1 - \text{PPV}$$

[false omission rate](#) (FOR)

$$\text{FOR} = \frac{\text{FN}}{\text{FN} + \text{TN}} = 1 - \text{NPV}$$



behzad.amanpour