Machine Learning 5

Evaluation Metrics

Accuracy – Sensitivity – Specificity
Precision – F1 Score – Confusion Matrix

behzad.amanpour

Accuracy

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

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

Behzad Amanpour

Sensitivity - Specificity

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

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

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

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

Behzad Amanpour

Sensitivity - Specificity

- **Precision, Positive Predictive Value**

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

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

$$\mathrm{F}_1 = 2 \times \frac{\mathrm{PPV} \times \mathrm{TPR}}{\mathrm{PPV} + \mathrm{TPR}} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}$$

Behzad Amanpour

Example

20 samples: 10 cancer + 10 normal

- **Algorithm1:**

  - all predicted cancer

$$TPR = \frac{TP}{P} = 100\% \qquad PPV = \frac{TP}{TP + FP} = 50\% \qquad TNR = \frac{TN}{N} = 0\%$$

- **Algorithm2:**

  - 5 cancerous samples predicted correctly

  - 15 samples predicted normal

$$TPR = \frac{TP}{P} = 50\% \qquad PPV = \frac{TP}{TP + FP} = 100\% \qquad TNR = \frac{TN}{N} = 100\%$$

Behzad Amanpour

Wikipedia

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

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. E

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 wh
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)

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

specificity, selectivity or true negative rate (TNR)

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

precision or positive predictive value (PPV)

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

negative predictive value (NPV)

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

miss rate or false negative rate (FNR)

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

fall-out or false positive rate (FPR)

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

false discovery rate (FDR)

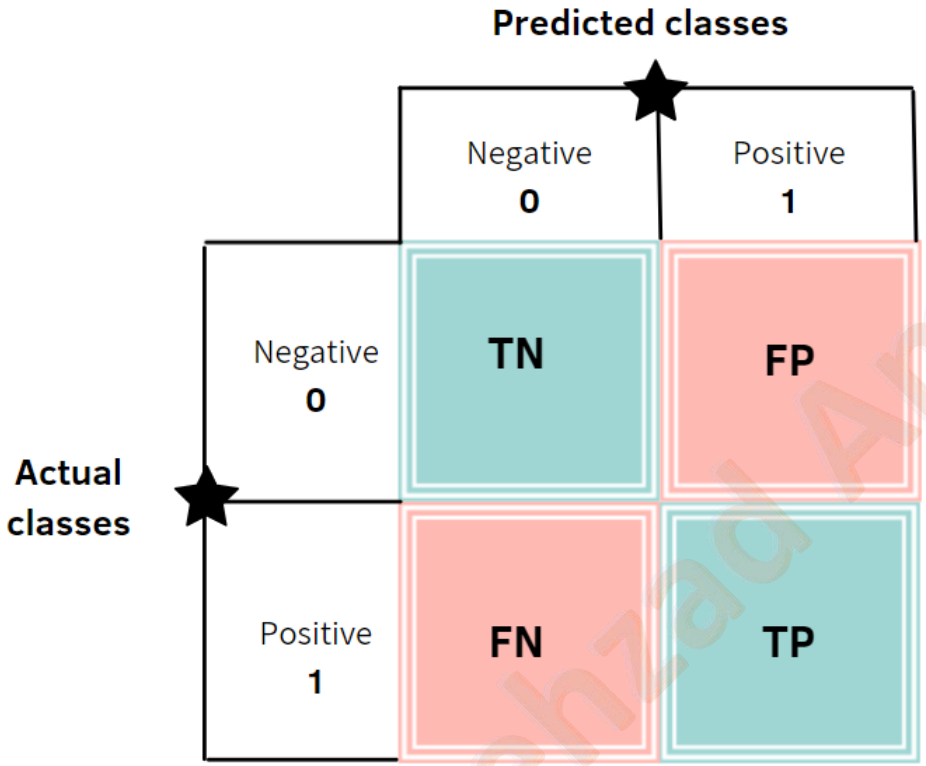
$$\mathrm{FDR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TP}} = 1 - \mathrm{PPV}$$

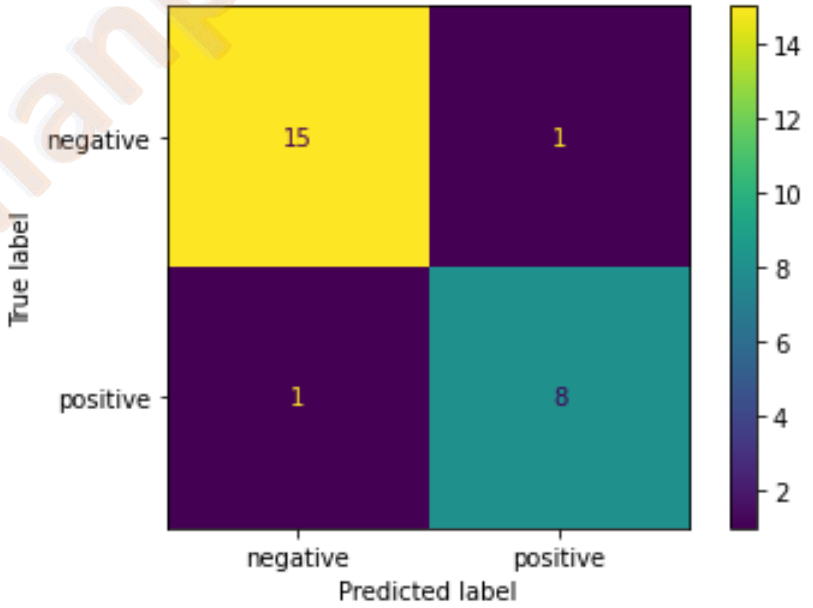
false omission rate (FOR)

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

Behzad Amanpour

Confusion Matrix (Matching Matrix)





https://towardsai.net/p/data-science/how-to-evaluate-you-model-using-the-confusion-matrixe

Behzad Amanpour