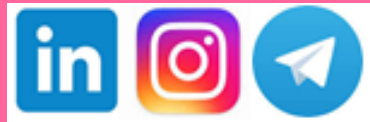


Machine Learning Evaluation Metrics Confusion Matrix ROC Curve



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, Recall, Selectivity, True Negative Rate**

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



Precision – F1 Score

- **Precision, Positive/Negative Predictive Value**

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

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

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



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**

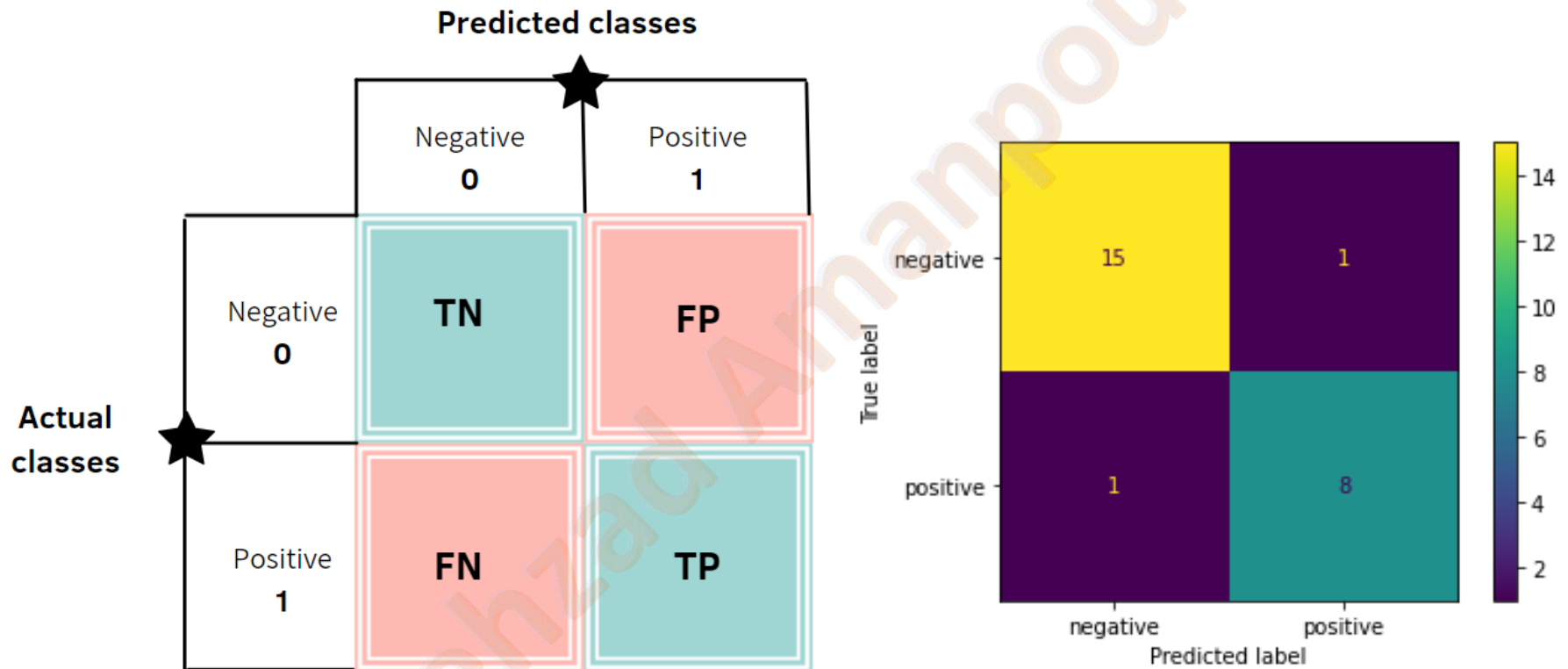
$$\text{TPR} = \frac{\text{TP}}{\text{P}} = 50\%$$

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

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



Confusion Matrix (Matching Matrix)

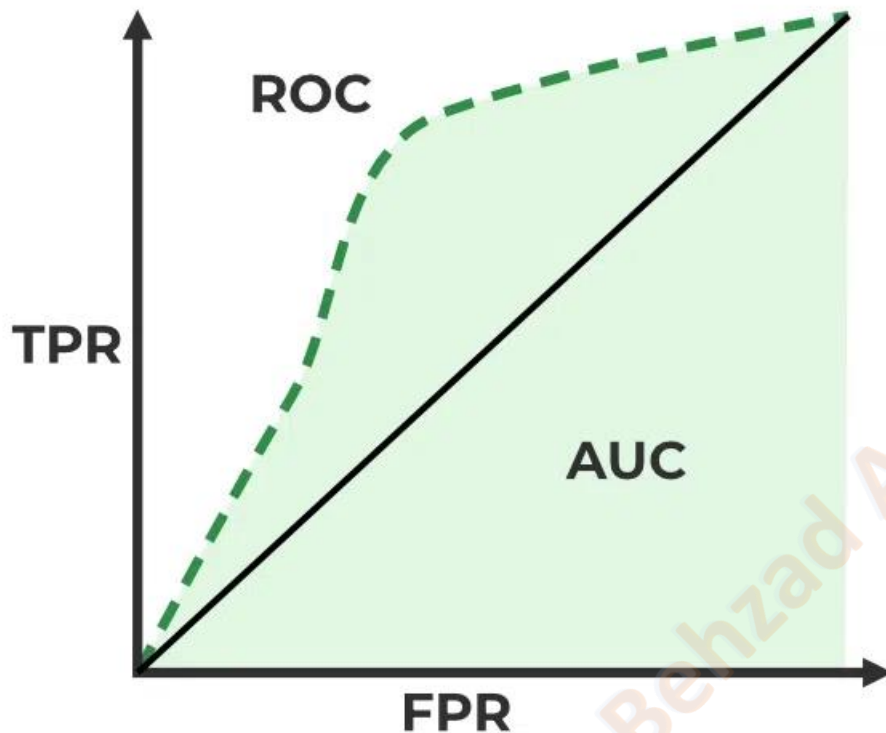


<https://towardsai.net/p/data-science/how-to-evaluate-your-model-using-the-confusion-matrix>



ROC Curve

ROC curve is obtained by plotting TPR against FPR at different **discrimination thresholds**.



<https://www.geeksforgeeks.org/auc-roc-curve/>



Logistic Regression Discrimination Threshold

- Logistic regression model predicts $P(Y=1)$ as a function of X .

$$\text{Ln}\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

<https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8>
<https://www.theanalysisfactor.com/what-is-logit-function/>



ROC Curve

