Understanding Machine Learning Metrics

Evaluating the performance of a machine learning model is crucial to ensuring its reliability in real-world applications. This article explores essential classification and regression metrics, their formulas, and visual explanations using probability distributions.

Classification Metrics ¶

1. The Confusion Matrix

The confusion matrix is the foundation for most classification metrics. It categorizes predictions into four types:

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Using this matrix, we define key classification metrics.

2. Accuracy

Formula:

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

Accuracy measures the proportion of correctly classified instances. While useful, it may be misleading in imbalanced datasets where one class dominates.

3. Precision (Positive Predictive Value - PPV)

Formula:

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

Precision answers the question: "Of all the predicted positives, how many were actually positive?" It is crucial in applications where false positives are costly, such as fraud detection.

4. Recall (Sensitivity, True Positive Rate - TPR)

Formula:

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

Recall measures the ability of a model to identify positive cases. High recall is essential in applications where missing a positive instance is critical, such as disease diagnosis.

5. Specificity (True Negative Rate - TNR)

Formula:

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

Specificity focuses on correctly identifying negative instances, useful in scenarios like spam detection.

6. Negative Predictive Value (NPV)

Formula:

$$NPV = \frac{TN}{TN + FN}$$

NPV measures the proportion of predicted negative instances that are truly negative.

7. F1 Score

Formula:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The F1 Score is the harmonic mean of precision and recall, providing a balanced measure when both are important.

8. ROC-AUC (Receiver Operating Characteristic - Area Under Curve)

The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR), showing the trade-off between sensitivity and specificity at different thresholds. AUC represents the area under this curve, where higher values indicate better performance.

9. PR-AUC (Precision-Recall Curve - Area Under Curve)

The PR curve plots precision against recall, useful for imbalanced datasets where ROC-AUC may not be informative.

Supervised Regression Metrics

1. Mean Absolute Error (MAE)

Formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAE calculates the average absolute difference between actual and predicted values. It provides a straightforward interpretation of error magnitude.

2. Mean Squared Error (MSE)

Formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

MSE penalizes larger errors more than MAE, making it sensitive to outliers.

3. Root Mean Squared Error (RMSE)

Formula:

$$RMSE = \sqrt{MSE}$$

RMSE represents the standard deviation of residuals, giving an error measurement in the same unit as the target variable.

4. R-Squared (R² Score)

Formula:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

R^2 measures how well the model explains variance in the target variable. Values close to 1 indicate better fit.

Metrics Sommary Table

Metric	Description	Formula	Used For
Confusion Matrix	A table summarizing the performance of a classification algorithm.	$\text{Confusion Matrix} = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$	Evaluating classification performance
Accuracy	The ratio of correct predictions to total predictions.	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$	General performance, classification
Precision (PPV)	The ratio of true positives to total predicted positives.	$\text{Precision} = \frac{TP}{TP+FP}$	Relevance of positive predictions
Recall (Sensitivity, TPR)	The ratio of true positives to total actual positives.	$\text{Recall} = \frac{TP}{TP+FN}$	Identifying all relevant instances
Specificity (TNR)	The ratio of true negatives to total actual negatives.	$\text{Specificity} = \frac{TN}{TN+FP}$	Evaluating performance in identifying negatives
Negative Predictive Value (NPV)	The ratio of true negatives to total predicted negatives.	$\text{NPV} = \frac{TN}{TN+FN}$	Identifying the relevance of negative predictions
F1 Score	The harmonic mean of precision and recall.	$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	Balance between precision and recall
Mean Absolute Error (MAE)	The average of absolute errors between predicted and actual values.	$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$, where y_i is the actual value and $\hat{y}_i$ is the predicted value.	Regression models
Mean Squared Error (MSE)	The average of the squared differences between predicted and actual values.	$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, where y_i is the actual value and $\hat{y}_i$ is the predicted value.	Regression models
Root Mean Squared Error (RMSE)	The square root of MSE, providing error in the original units.	$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	Regression models

Metric	Description	Formula	Used For
R-Squared (R^2 Score)	The proportion of variance in the dependent variable explained by the model.	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$, where y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the actual values.	Evaluating regression models

Visualizing the Metrics

1. Classification Decision Boundary

This plot illustrates the effect of a decision threshold on predictions:

- Blue area represents True Positives (TP) and False Negatives (FN).
- Red area represents False Positives (FP) and True Negatives (TN).
- The black dashed line marks the classification threshold.

Classification Decision Boundary

2. Sensitivity & Specificity

This plot highlights how well the model distinguishes between classes:

- True Positives (TP) in blue.
- True Negatives (TN) in gray.
- False Positives (FP) in red.
- False Negatives (FN) in green.

Sensitivity and Specificity

3. Precision (PPV) & Negative Predictive Value (NPV)

This visualization emphasizes the proportion of correctly predicted positives and negatives.

PPV and NPV

4. F1 Score Balance

The F1 Score balances precision and recall. This visualization demonstrates how a shift in the threshold affects these metrics.

F1 Score Balance

5. Regression Metrics Comparison

This chart compares MAE, MSE, and RMSE across different models and error distributions.

Regression Metrics Comparison

6. R-Squared (R^2) Visualization

This plot shows how well predicted values align with actual values, indicating model fit quality.

R-Squared Visualization

Conclusion

Choosing the right metric depends on the problem. For classification tasks, precision, recall, F1-score, and PR-AUC are more informative than accuracy in imbalanced datasets. ROC-AUC is useful for balanced classification problems, while the confusion matrix provides a detailed error analysis.

I've added references to the screenshots of the plots in the article. Let me know if you need modifications or additional visualizations.

In [ ]: