Model

Evaluation

8 Essential Metrics
YOU NEED TO KNOW!



• •

Model Evaluation: Metric 01/08

Accuracy

.

- Explanation: Accuracy measures the overall correctness of the model's predictions by calculating the ratio of correctly predicted instances to the total number of instances.
- Use Case: Accuracy is commonly used for classification problems when the class distribution is balanced.
- Range & Interpretation: The accuracy score ranges from 0 to 1, where 1 indicates a perfect prediction. However, accuracy alone may not be reliable if the dataset is imbalanced.
 - Formula:

 $(Number\,of\,Correct\,Predictions\,) \ (Total\,Number\,of\,Predictions\,)$





Precision

- Explanation: Precision calculates the proportion of correctly predicted positive instances out of all instances predicted as positive.
 - Use Case: Precision is useful when the focus is on minimizing false positives.
- Range & Interpretation: Precision ranges from 0 to 1,
 where 1 represents a perfect precision score. A higher
 precision indicates a lower rate of false positives.

• Formula:
$$\frac{\left(True\,Positives\,\right)}{\left(True\,Positives+False\,Positives\,\right)}$$



Model Evaluation: Metric 03/08

Recall

- Explanation: Recall, also known as sensitivity or True
 Positive Rate, measures the proportion of correctly predicted positive instances out of all actual positive instances.
 - Use Case: Recall is valuable when the goal is to minimize false negatives.
 - Range & Interpretation: Recall ranges from 0 to 1, with 1 indicating perfect recall. A higher recall implies a lower rate of false negatives.

• Formula:
$$\frac{\left(True\,Positives\,\right)}{\left(True\,Positives + False\,Negatives\,\right)}$$

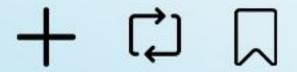




F1 Score

- Explanation: The F1 score combines precision and recall into a single metric, providing a balance between the two. It is the harmonic mean of precision and recall.
 - Use Case: F1 score is useful when both false positives and false negatives need to be considered equally.
- Range & Interpretation: The F1 score ranges from 0 to 1, with 1 being the best possible value. It represents the trade-off between precision and recall.

Formula:
$$2 \times \left(\frac{(Precision \times Recall)}{(Precision + Recall)}\right)$$





Model Evaluation: Metric 05/08

ROC-AUC

Receiver Operating Characteristic - Area Under the Curve

- **Explanation:** ROC-AUC measures the performance of a binary classifier by analyzing the relationship between the TPR and the FPR at different threshold settings.
- Use Case: ROC-AUC is suitable for imbalanced datasets, provides a comprehensive evaluation of the classifier's performance.
 - Range & Interpretation: The ROC-AUC score ranges from 0 to 1, where 1 indicates a perfect classifier.
 A higher ROC-AUC score implies better model performance.

· Formula:
$$\int\limits_0^1 TPR(FPR)dFPR$$

TPR: True Positive Rate FPR: False Positive Rate



Swipe to Learn



Model Evaluation: Metric 06/08



- Explanation: MSE calculates the average squared difference between predicted and actual values, providing a measure of the model's performance for regression problems.
 - Use Case: MSE is commonly used when the emphasis is on large errors.
- Range & Interpretation: The MSE value can range from 0 to positive infinity. A lower MSE indicates better model performance, with 0 being the ideal value.

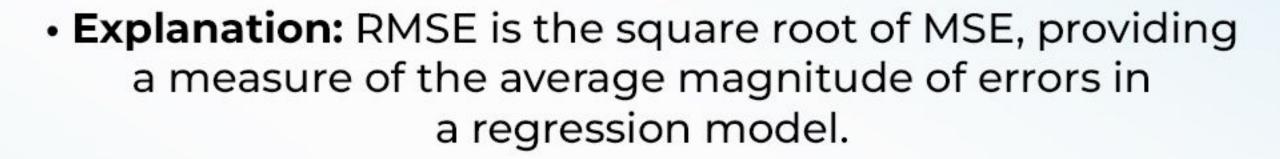
$$\left(\frac{1}{n}\right) \times \sum_{true} \left(y_{true} - y_{pred}\right)^2$$



Model Evaluation: Metric 07/08



Root Mean Squared Error



- Use Case: RMSE is useful when the emphasis is on both small and large errors.
- Range & Interpretation: RMSE values range from 0 to positive infinity. A lower RMSE indicates better model performance, with 0 being the ideal value.

• Formula:
$$\sqrt{MSE} = \sqrt{\left(\frac{1}{n}\right)} \times \sum_{n=1}^{\infty} \left(y_{true} - y_{pred}\right)^2$$



Model Evaluation: Metric 08/08



- Explanation: R-squared measures the proportion of the response variable's variance that is captured by the model.
 It indicates how well the regression model fits the data.
 - Use Case: R-squared is commonly used to assess the goodness of fit in regression models.
 - Range & Interpretation: R-squared values range from 0 to 1, where 1 represents a perfect fit. A higher R-squared indicates better model performance.

$$1 - \frac{RSS}{TSS}$$

RSS: Residual Sum of Squares

TSS: Total Sum of Squares

