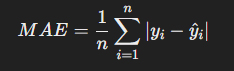
**Performance Metrics**

**Regression**

**Mean Absolute Error (MAE)**

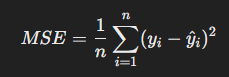


Purpose: Measures the average magnitude of errors between predictions and actual.

Pros: Simple to interpret; doesn’t heavily penalize large errors.

Cons: Doesn’t capture variability very well; less sensitive to outliers.

**Mean Squared Error (MSE)**

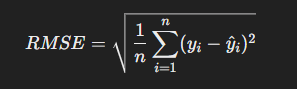
****

Purpose: Measures average squared errors, penalizing large errors more than MAE.

Pros: Useful for algorithms that need error gradients (in optimization).

Cons: Sensitive to outliers, can disproportionately affect the metric.

**Root Mean Squared Error (RMSE)**

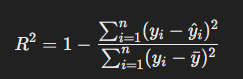
****

Purpose: Square root of MSE, providing errors in same units as target variable.

Pros: Benefits of MSE but is more interpretable.

Cons: Sensitivity to outliers, as in MSE.

**R-Squared (R2)**

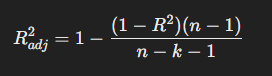
****

Purpose: Indicates the proportion of variance in the dependent variable explained by the model.

Pros: Easy to interpret, general indicator of model performance.

Cons: Can be misleading for small datasets or overfitting, doesn’t penalize complexity.

**Adjusted R-Squared**

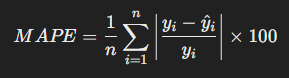
****

Purpose: Similar to R2, adjusts for the number of predictors, penalizing overfitting.

Pros: Accounts for model complexity; useful in feature selections/

Cons: Less intuitive to explain than R2.

**Mean Absolute Percentage Error (MAPE)**

****

Purpose: Expresses errors as a percentage of actual values.

Pros: Intuitive for business applications and relative comparisons.

Cons: Undefined when actual values are zero; biased towards underestimating errors.

**Akaike Information Criterion/Bayesian Information Criterion**

** **

Purpose: Penalize model complexity while considering goodness of fit. Often used to compare separate models.

Pros: Useful for model comparison.

Cons: Not a standalone metric.

**Classification**

**Accuracy**

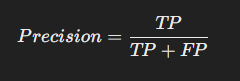
****

Purpose: Measures the proportion of correct predictions.

Pros: Easy to understand; broad indicator of performance.

Cons: Misleading for imbalanced datasets (high accuracy with majority class decision).

**Precision**

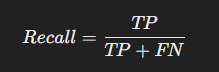
****

Purpose: Proportion of true positives among predicted positives.

Pros: Useful when false positives are costly.

Cons: Doesn’t account for false negatives.

**Recall (Sensitivity)**

****

Purpose: Proportion of true positives identified out of all true positives.

Pros: Useful when false negatives are costly.

Cons: Doesn’t account for false positives.

**F1 Score**

****

Purpose: Harmonic mean of precision and recall.

Pros: Balances precision and recall, especially useful for imbalanced datasets.

Cons: Doesn’t reflect true negatives; not as intuitive as individual metrics.

**Specificity**

****

Purpose: Proportion of true negatives among actual negatives.

Pros: Complements recall; useful in medical diagnostics.

Cons: Not always emphasized in standard evaluations.

**ROC/AUC**

****

Purpose: Plots true positive rate vs false positive rate at various thresholds.

Pros: Visualizes trade-offs between sensitivity and specificity.