Cluster Evaluation Metrics

📊 Internal Metrics

🎯 Silhouette Score

`Range: [-1, 1]`

Measures separation and cohesion

Higher is better

📉 Davies-Bouldin Index

`Range: [0, ∞)`

Ratio of within to between cluster distances

Lower is better

📈 Calinski-Harabasz Index

`Range: [0, ∞)`

Ratio of between to within variance

Higher is better

🎲 Inertia (WCSS)

`Range: [0, ∞)`

Within-cluster sum of squares

🏷️ External Metrics

When ground truth labels are available

🖇️ Adjusted Rand Index (ARI)

`Range: [-1, 1]`

Similarity between two clusterings

Closer to 1 is better

ℹ️ Normalized Mutual Info (NMI)

`Range: [0, 1]`

Mutual information between clusterings

Closer to 1 is better

📊 Fowlkes-Mallows Index

`Range: [0, 1]`

Geometric mean of precision and recall

Closer to 1 is better

Minimize for K-means

Visual Methods

Silhouette plots, cluster distributions

Domain Validation

Do clusters make business sense?

Domain validation is the most important!

Try Interactive Clustering Visualization

Test K-means, DBSCAN, Hierarchical, and MeanShift clustering algorithms

Open Clustering Visualizer →