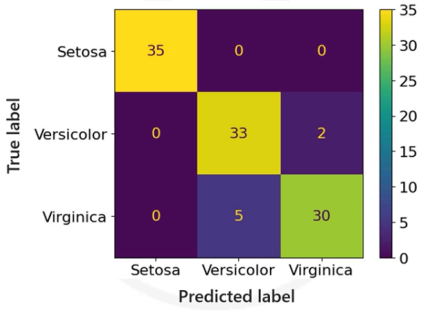
Module 5

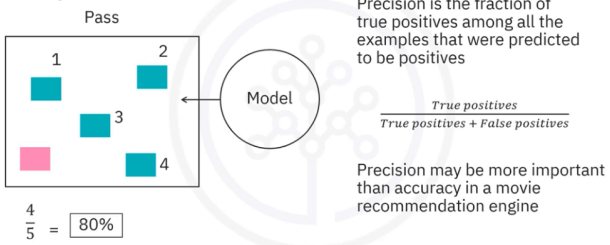
Thursday, July 03, 2025 10:49 AM

Iris confusion matrix



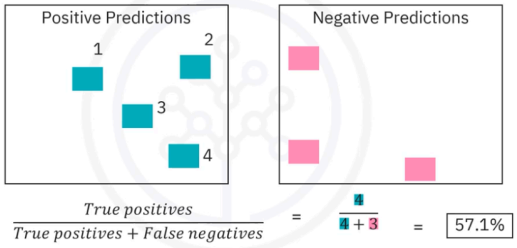
Precision

Example



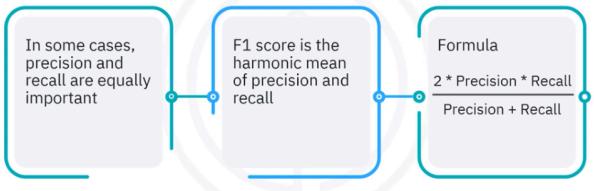
Recall

Example



F1 score

Example



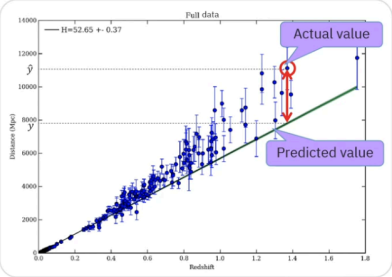
F1 score

	Precision	Recall	F1 score	Support
Setosa	1.00	1.00	1.00	35
Versicolor	0.87	0.94	0.90	35
Virginica	0.94	0.86	0.90	35
Accuracy	-	-	0.93	105
Weighted avg	0.94	0.93	0.93	105

Evaluating regression models

- Regression models make prediction errors
- **Evaluating a regression model**: Determining how accurately it can predict continuous numerical values

Error of the model



Error: Measure of the difference between the data points and the trend line generated by the algorithm

Regression metrics

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

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

$$\text{RMSE} = \sqrt{\text{MSE}}$$

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

Provide insight into a model's performance

- Accuracy
- Error distribution
- Error magnitude

Explained variance and R²

$$\text{Explained Variance} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$

$$\text{Unexplained Variance} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \text{n} * \text{MSE}$$

$$\text{Total Variance} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sigma^2$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\text{Unexplained Variance}}{\text{Total Variance}}$$

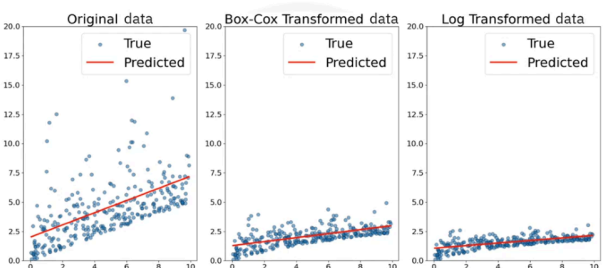
Explained variance and R²

Perfect predictor
If $(\hat{y}_i - y_i)$ for all i :
- Explained variance = Total variance
- Unexplained variance = 0
- $R^2 = 1$

Negative R²
Unexplained variance > Total variance

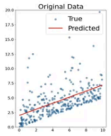
Mean-value model
If $(\hat{y}_i - \bar{y})$ for all i:
- Explained variance = 0
- Total variance = Unexplained variance
- $R^2 = 0$

Regression metric comparison

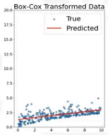


Regression metric comparison

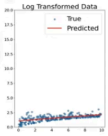
Untransformed data
Explained variance: 0.30
R^2: 0.29
MAE: 1.54
MSE: 4.767
RMSE: 2.183



Box-cox-transformed data
Explained variance: 0.40
R^2: 0.40
MAE: 0.47
MSE: 0.366
RMSE: 0.605



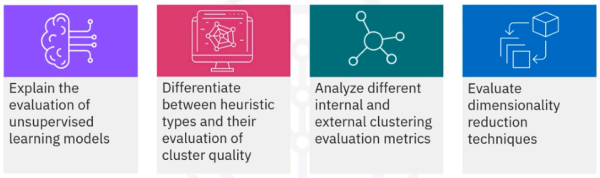
Log-transformed data
Explained variance: 0.42
R^2: 0.42
MAE: 0.29
MSE: 0.148
RMSE: 0.375



Recap

- **Evaluating a regression model**: Determine how accurately it can predict continuous numerical values
- **Model error**: Measure of the difference between the data points and the trend line
- **Essential regression metrics**: MAE, MSE, RMSE, and R2
- **Explained variance**: Sum of squared differences between the predictions and the average value of the actual target data
- **R2**: Measures the proportion of variance in target variable predictable from input variables

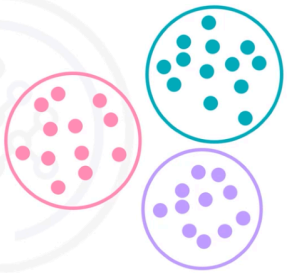
What you will learn



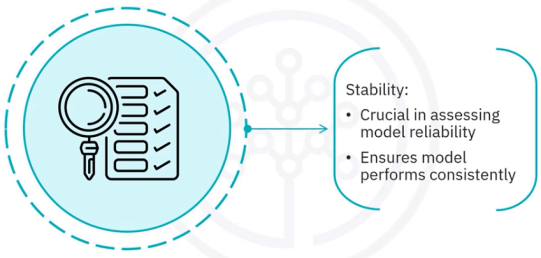
What is unsupervised learning evaluation?

Clustering and dimensionality reduction:

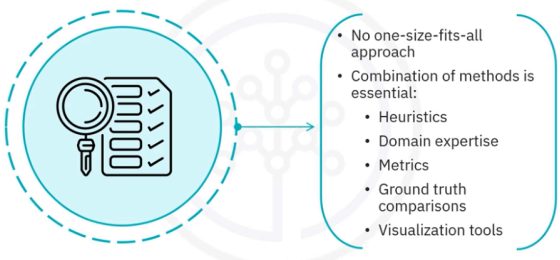
- Focus on discovering hidden patterns and structures
- Assess pattern quality for model effectiveness



Why is evaluation critical?



Why is evaluation critical?



Important clustering heuristics

- Group similar data points into clusters
- Unsupervised learning lacks predefined labels
- Heuristics are used to evaluate cluster quality

Important clustering heuristics

Internal clustering evaluation metrics

- Assess the quality of clustering
- Focuses on input data

Internal clustering evaluation metrics

Silhouette score:

- Compares cohesion within each cluster
- Assesses separation from other clusters
- Ranges from -1 to 1
- Higher values indicate better-defined clusters

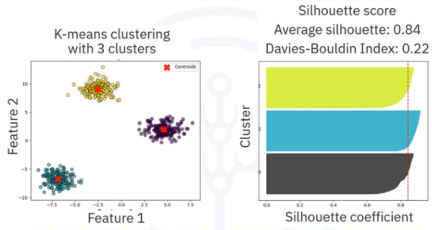
Davies-Bouldin Index:

- Measures cluster compactness ratio
- Assesses separation from the nearest cluster
- Lower values indicate more distinct clusters
- Indicates more compact clusters

Inertia:

- Used in k-means clustering
- Calculates the sum of variances within clusters
- Lower values suggest more compact clusters
- Increasing clusters reduces variance
- Creates a trade-off in clustering

Internal clustering evaluation



- Applies k-means to simulated blobs
- Obtains distinct and dense clusters

Recap

- Explain the evaluation of unsupervised learning models
- Explain unsupervised learning results and stability
- Differentiate between the different types of heuristics
- Analyze different internal clustering evaluation metrics
- Evaluate internal clustering by applying k-means
- Analyze external clustering evaluation metrics
- Evaluate dimensionality reduction

Your grade: 100%

Your latest: **100%•**
Your highest: **100%•**
To pass you need at least 60%. We keep your highest score.

1.
Question 1

What is the purpose of a train/test split in machine learning?

To increase the size of the training dataset

To optimize the model's hyperparameters

To visualize the dataset

To estimate the performance of machine learning algorithms on unseen data

Correct

This approach helps evaluate the model's ability to generalize to data it hasn't seen before, giving an estimate of real-world performance.

1 / 1 point

2.
Question 2

What does the F1 score represent in model evaluation?

The harmonic mean of precision and recall

The weighted sum of precision and recall

The probability of a true positive prediction

The average accuracy of the model

Correct

F1 score calculates a balanced mean between precision and recall, making it valuable when both metrics need to be optimized simultaneously.

1 / 1 point

3.
Question 3

What does R-squared measure in regression analysis?

The unexplained variance of the model

The proportion of variance in the target variable explained by the model

The total variance of the target variable

The sum of squared differences between predicted and actual values

Correct

R-squared represents the proportion of variance in the target variable that can be predicted by the input features, indicating model performance.

1 / 1 point

4.
Question 4

What is the effect of using a mean-value model in regression analysis?

R-squared will be one half

R-squared will be one

R-squared will be zero, as the model explains no variance

Explained variance will be greater than total variance

Correct

A mean-value model predicts the average value for all data points, resulting in an R-squared value of zero, indicating no explanatory power.

1 / 1 point

5.
Question 5

Which of the following clustering evaluation metrics ranges from -1 to 1, with higher values indicating better-defined clusters?

Silhouette Score

Inertia

Davies-Bouldin Index

Within-cluster sum of squares (WCSS)

Correct

The Silhouette Score measures cluster cohesion and separation, with higher values indicating more defined and distinct clusters.