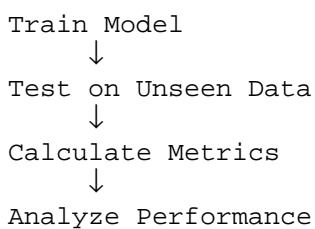# Model Evaluation

Introduction

Model evaluation is the process of assessing how well a machine learning model performs on unseen data. It helps determine whether the model is accurate, reliable, and suitable for real-world use.

Why Model Evaluation is Important

1 Measures model performance objectively

2 Helps compare different models

3 Detects overfitting and underfitting

4 Ensures generalization to new data

Model Evaluation Workflow

```
Train Model
     ↓
Test on Unseen Data
     ↓
Calculate Metrics
     ↓
Analyze Performance
```

Evaluation Metrics for Classification

1 **Accuracy:** Correct predictions / Total predictions

2 **Precision:** Correct positive predictions / Total predicted positives

3 **Recall:** Correct positive predictions / Total actual positives

4 **F1-score:** Harmonic mean of precision and recall

Confusion Matrix

```
Predicted
        Positive   Negative
Actual  TP         FN
        FP         TN
```

Evaluation Metrics for Regression

1 Mean Absolute Error (MAE)

2 Mean Squared Error (MSE)

3 Root Mean Squared Error (RMSE)

4 R-squared (R²)

Train-Test Split

Model evaluation typically uses a train-test split to evaluate performance on unseen data.

```
Dataset
  ██ Training Set (80%)
  ██ Test Set (20%)
```

Cross Validation

Cross-validation improves evaluation reliability by splitting data into multiple folds and training/testing the model multiple times.

```
Data → Split into K folds → Train/Test K times → Average Performance
```

Advantages of Proper Model Evaluation

1   Better model selection

2   Improved reliability

3   Prevents misleading results

Real-Life Example

In credit risk analysis, model evaluation ensures that loan approval models accurately predict default risks on new customers.

Summary

Model evaluation ensures that machine learning models perform well on unseen data. Using appropriate metrics and validation techniques is essential for building trustworthy ML systems.