# Cross-validation Strategies

## K-Fold CV

| Test | Train | Train | Train | Train |
|------|-------|-------|-------|-------|

Rotate
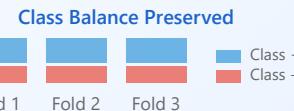
Split data into K folds, train on K-1, test on 1, repeat K times

**Standard: K=5 or 10**

## Stratified CV

**Class Balance Preserved**

Fold 1  Fold 2  Fold 3

Class +
Class -

Preserve class distribution in each fold - crucial for imbalanced data

**Use for classification**

## Leave-One-Out

Test          Train N-1
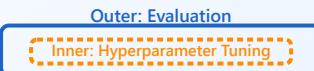
*Repeat N times*

Train on N-1 samples, test on 1, repeat N times

**For very small N**

## Nested CV

**Outer: Evaluation**

Inner: Hyperparameter Tuning
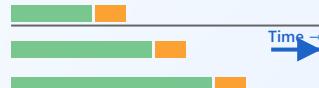
Unbiased performance estimate

Outer loop for evaluation, inner loop for hyperparameter tuning

**Unbiased estimates**

## Time Series Split

Time →

Train on past, test on future - respects temporal order

**For longitudinal data**

## Group CV

P1   P2   P3

**All samples from same patient together!**

Keep all samples from same patient/site together

**Prevent data leakage**

---

⚠️ **Biomedical Pitfall**

Multiple samples from same patient must stay in same fold to avoid optimistic bias!