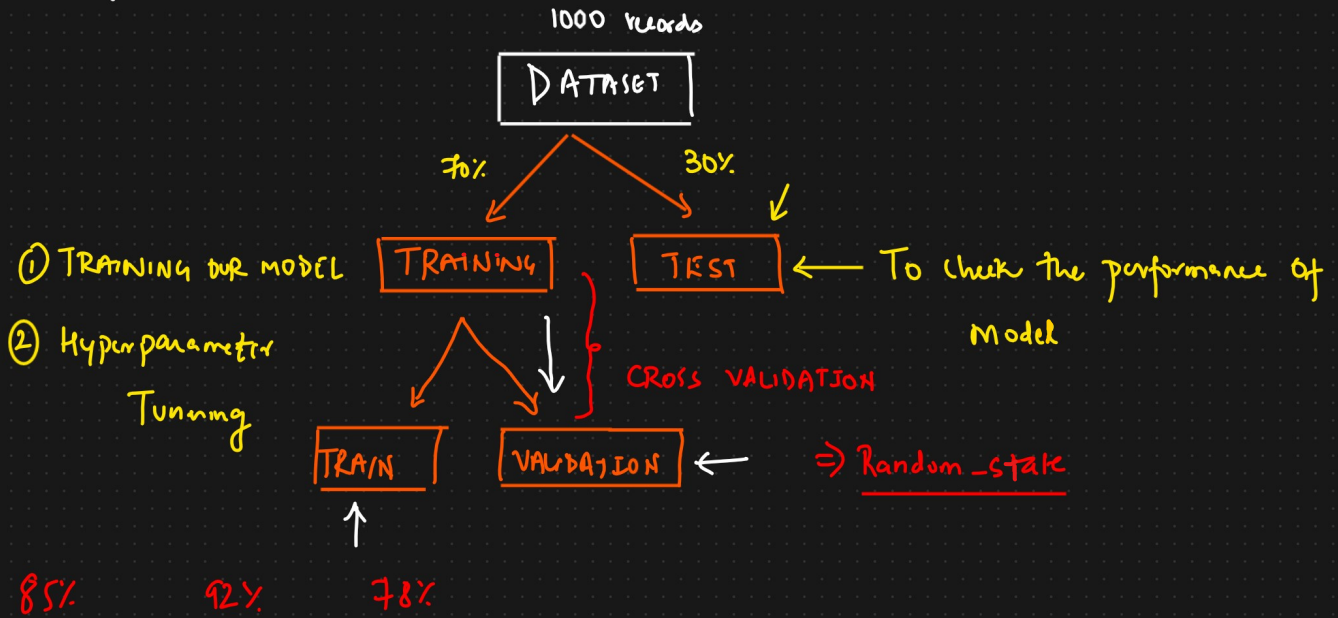


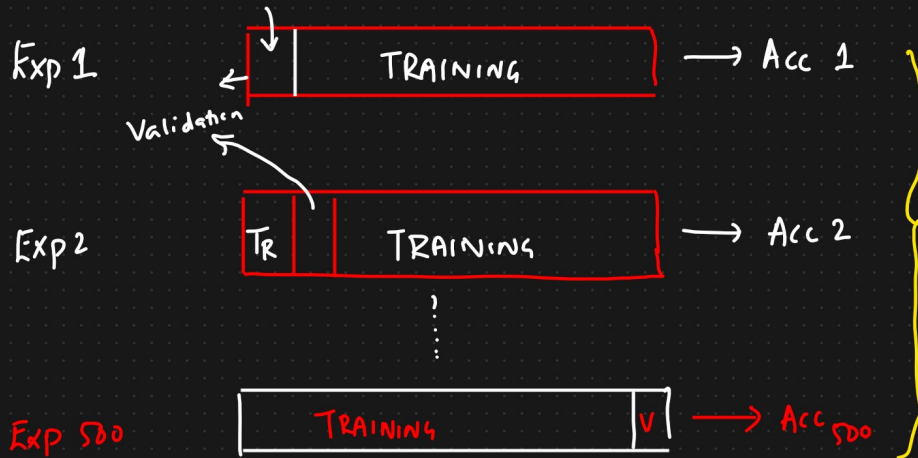
Types of CROSS VALIDATION

- Cross validation is a technique used in machine learning to evaluate the performance of a model on unseen data.



- ① leave One Out CV (LOOCV) ② leave p out CV
- $P=10, P=20, P=30$

TRAINING → 500 Records ↑↑ **Complexity of Training Model**



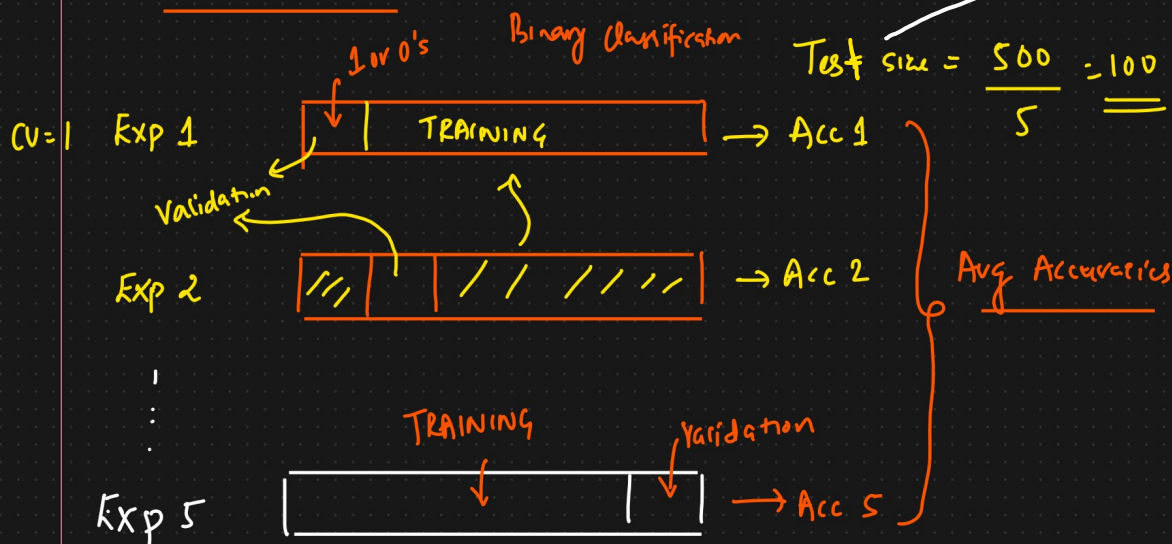
- ① **Overfitting** → Taking entire **TRAINING Data** → **Accuracy High** → New Test Data → **Acc ↓↓**
- Validation ↓↓ Data Acc ↓↓

③ K Fold CV

$K=5$

$n=500$

Validation Size

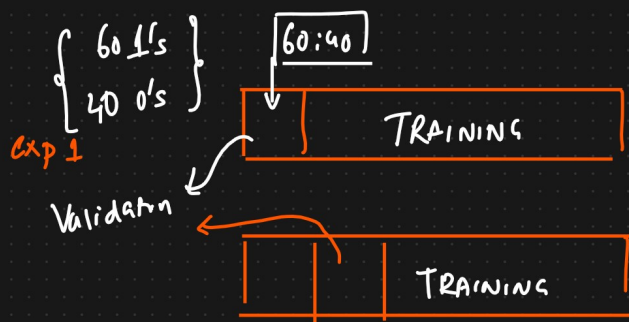


④ Stratified K Fold CV

Similar to K-Fold, but ensures that the ratio in validation data set fold matches the original dataset.

$K=5$

- Useful for imbalanced datasets



Test size = $\frac{500}{5} = 100$

↓

Validation Data

⑤ Time Series CV

- Used for time-dependent data,

Reviews

Product Sentiment Analysis

Time

JAN → DEC

TRAINING

Validation

DAY 1 DAY 2 DAY 3 DAY 4 | - - - DAY N

Time Series Application