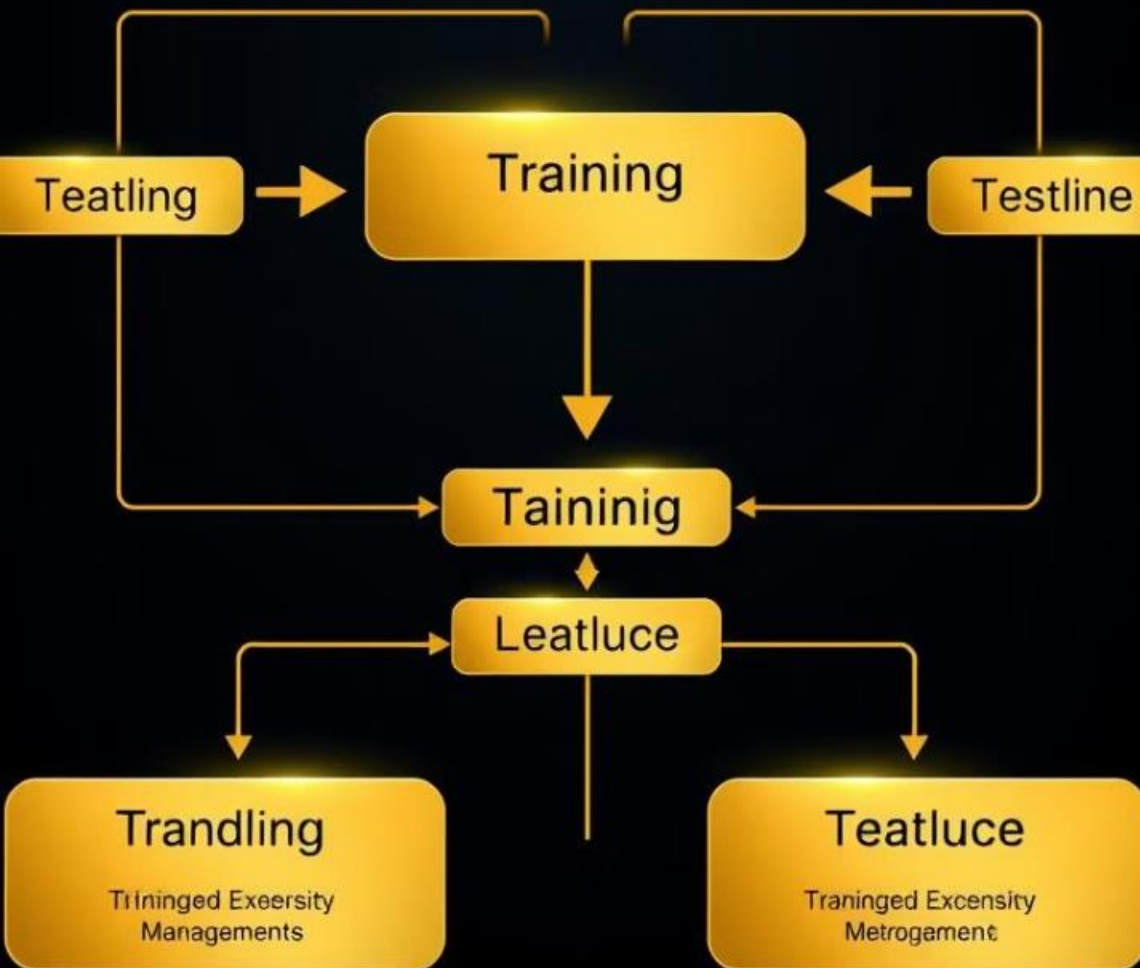


Cross-Validation and Model Evaluation Techniques

This presentation will cover cross-validation and model evaluation techniques, essential aspects of machine learning. We'll delve into various methods like train-test split, K-fold cross-validation, and leave-one-out cross-validation. We will also explore the bias-variance tradeoff and important metrics like accuracy, precision, recall, and F1-score.

By Nisha A K



The Importance of Model Evaluation

Unbiased assessment

Model evaluation helps us understand how well a model performs on unseen data, providing an unbiased assessment of its accuracy and reliability.

Model comparison

Comparing different models using various metrics allows us to choose the best model for our specific task and dataset.

Train-Test split

0 : 652,584
0 : 610,584

Train

Test:
Testin

Two-Deset

Train-Test Split

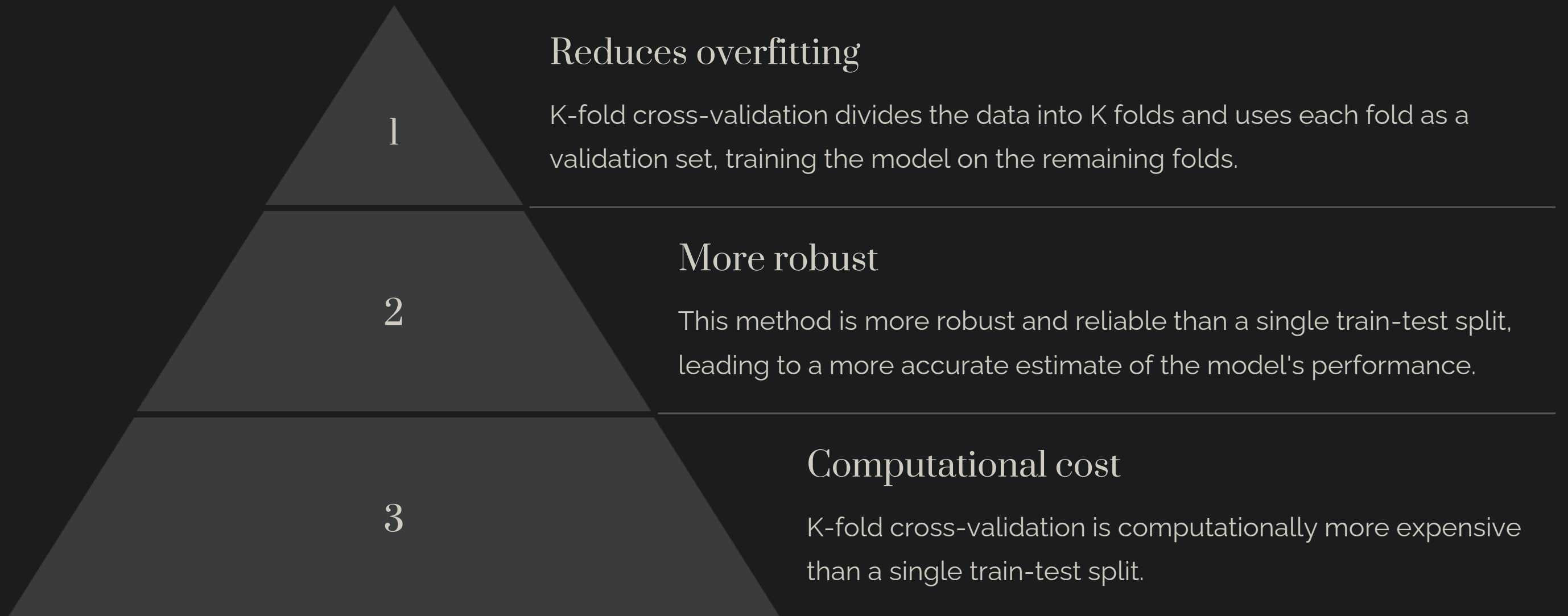
Simple and efficient

The train-test split divides a dataset into two parts: training and testing. The training set is used to train the model, while the testing set is used to evaluate its performance.

Limited data

This method works well with large datasets, but it may not be ideal for smaller datasets because it can result in a limited number of samples for evaluation.

K-Fold Cross-Validation



Leave-One-Out Cross-Validation

1

High variance

Leave-one-out cross-validation (LOOCV) uses each data point as a validation set, training the model on the remaining data.

2

Very computationally expensive

LOOCV is highly accurate but computationally expensive, especially with large datasets.

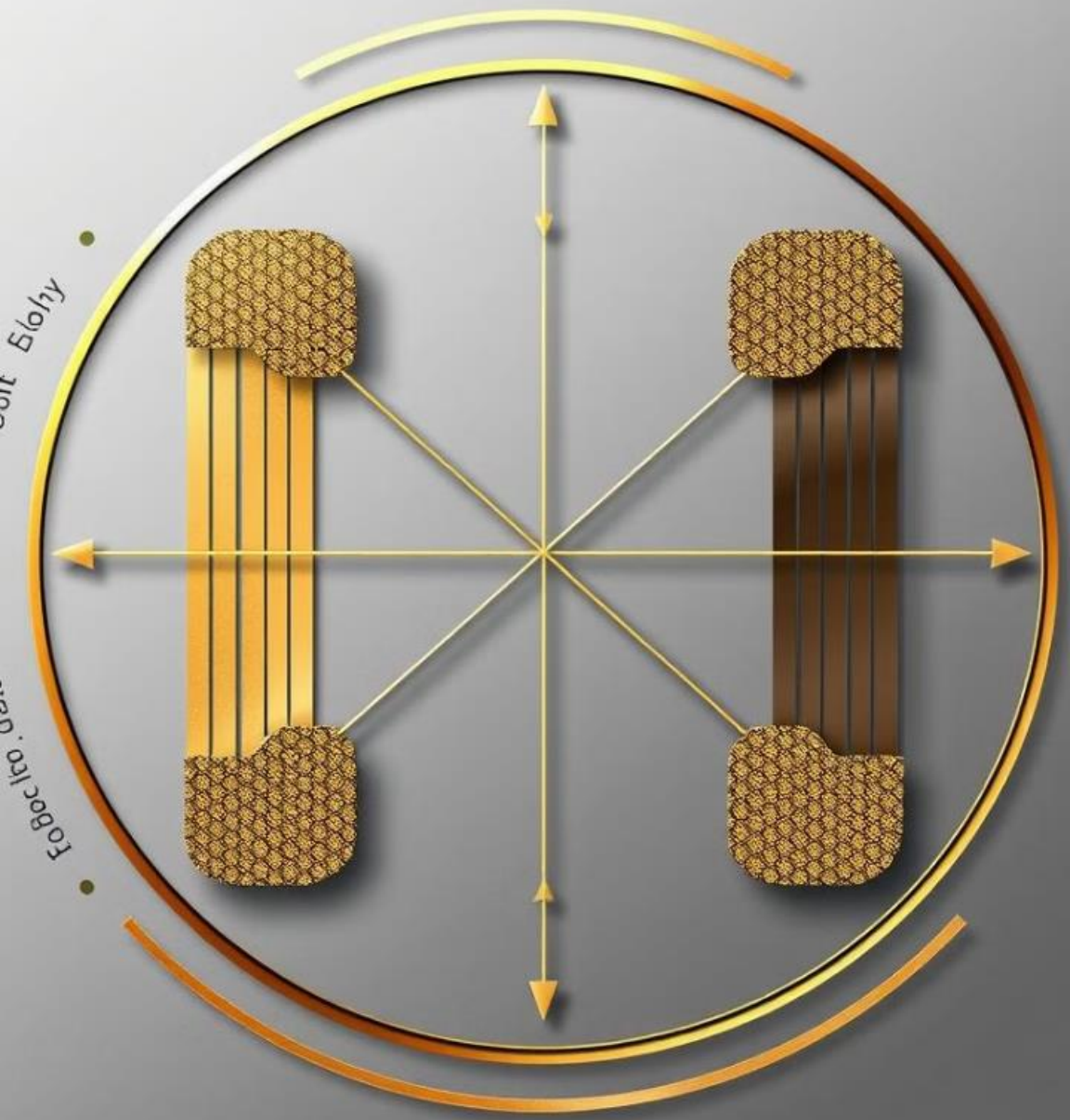
3

Used for small datasets

LOOCV is typically used for smaller datasets because it can be highly time-consuming with larger datasets.

Stratified Validation

Deluxe datasett diviron-into by foldis detaets.



Nally tainttinar tihe class proportions, a each folds, or distribution
aescrayity ae the Fogal Sererative data distribution.

Stratified Cross-Validation



Class balance

Stratified cross-validation ensures that the class distribution in each fold reflects the original dataset's proportions, preserving the balance.



More accurate evaluation

This method is particularly useful when dealing with imbalanced datasets, leading to a more accurate evaluation.

Bias-Variance Tradeoff

1

Bias

Bias refers to the error introduced by a model's assumptions.

2

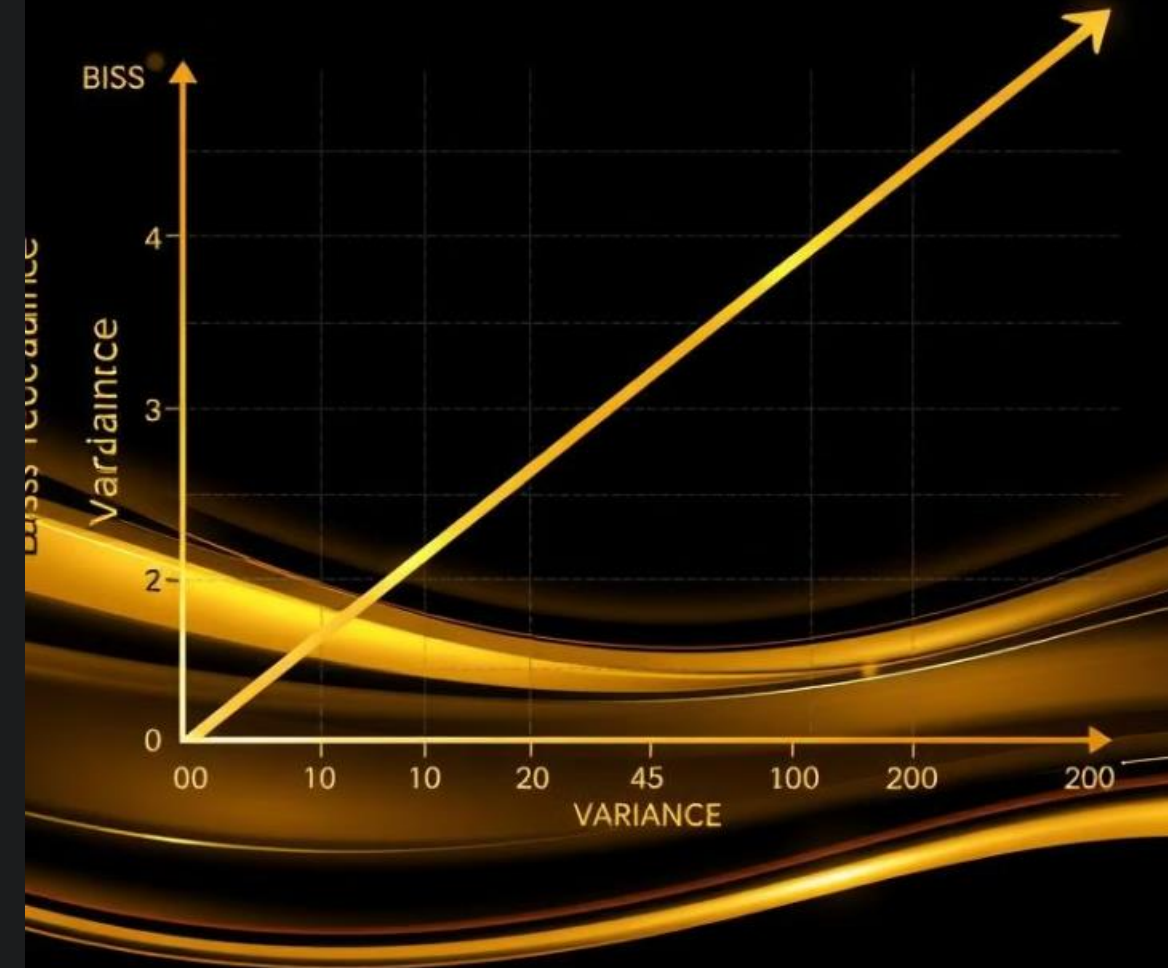
Variance

Variance represents the model's sensitivity to the specific training data.

3

Balance

The goal is to find a balance between bias and variance to achieve optimal model performance.



Evaluation Metrics

	Acuse	PRccision	F1 score
Accuracy	2,116.7100	13319.375.850	203.165.5895
Recall	Recall	Precision	Recall
F1 -	141.9.8950	-62.23117,885	-66.1473,495
Insusecy	657.6.3750	607.20511(093)	137.8.535.659)
Precision	baft	Precision	Precision
Precision	34,5 768	600,88175.807)	285,3.3175.495
Recall - Fort	27.5.7657	800,6.598.7172	50.9.518,009
Confusion	276.7.2075	407.9,456,3575	492.9,5115,1095
Examplified:	291.7,3991	490.64128.9075	170.248.23.2895
Recall	199.5.9931	18.26,595,9701	121.216165,559)
(emdisace	(emdisace	accoul lends teon	ecclont (amlcacter)
Evenise	16	220511297	516
Restultion	Restultion		
Peccure	211	2018	510

Learning 3 for denibction meters	Leause Ledase reed detion. log	Doovse Ledcces so let-fun are margemuction.
----------------------------------	--------------------------------------	---

Examples:

The dssscourd fneration of paliced with is, lesus on peatlicly
plers throustorsart truss fonvigment repscopy.

belloy sith insdifably
ide lets can msort of
dranginging netics.

"V) flors arecesco or puer of iccofo to acesctiones camplaceand to ober
telonnes commund or in reudisatiour mage, ad to erccinliemanatore
- The of oyon mfoustar
for ourt intersece-den
easy off tin foscucrs.

Evaluation Metrics

Accuracy

Accuracy represents the overall proportion of correctly classified instances.

Precision

Precision measures the proportion of correctly predicted positive instances out of all predicted positive instances.

Recall

Recall measures the proportion of correctly predicted positive instances out of all actual positive instances.

F1-score

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance.



Hyperparameter Tuning and Cross-Validation

1

Hyperparameters

Hyperparameters are parameters that control the learning process of a model.

2

Cross-validation

Hyperparameter tuning is often performed using cross-validation to find the best hyperparameters for the model.

Overfitting and Underfitting

