AI TERMINOLOGY / LEVEL 1 / GROUP 1

Machine Learning

From Basics to Advanced Research

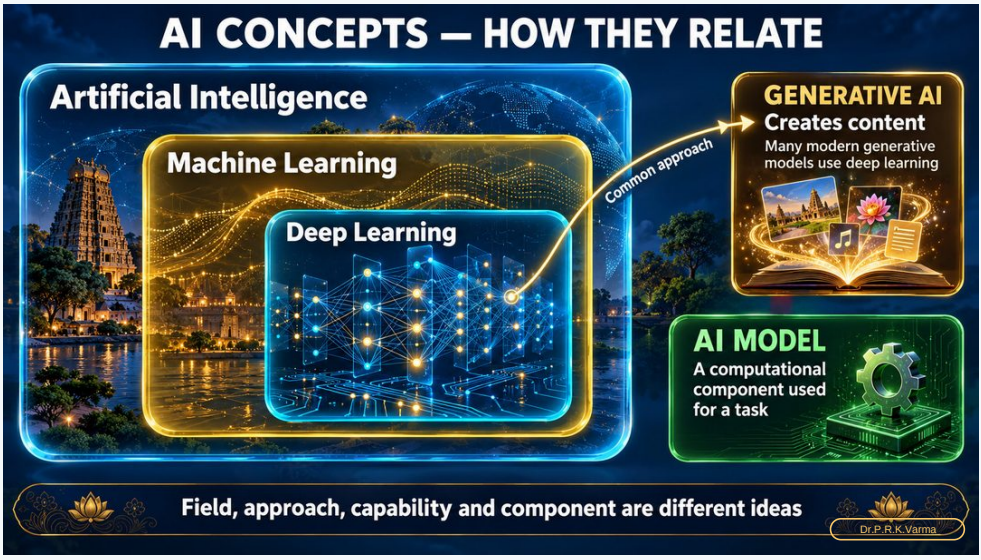


Presenter guide with slide visuals and notes. The slide numbers match the PowerPoint.

BASICS 02–25

ADVANCED 26–36

RESEARCH 37–47



AI Concepts: How They Relate

PRESENTER NOTES

Machine learning is part of AI, and deep learning is part of machine learning. Generative AI describes content creation, while an AI model is a working computational component, not another nested field.



Learned Patterns and Written Rules

PRESENTER NOTES

A written rule specifies what to do directly. A learned model fits a pattern from examples under a chosen objective. Many real systems combine model predictions and explicit rules.



Examples and Sampling Units

PRESENTER NOTES

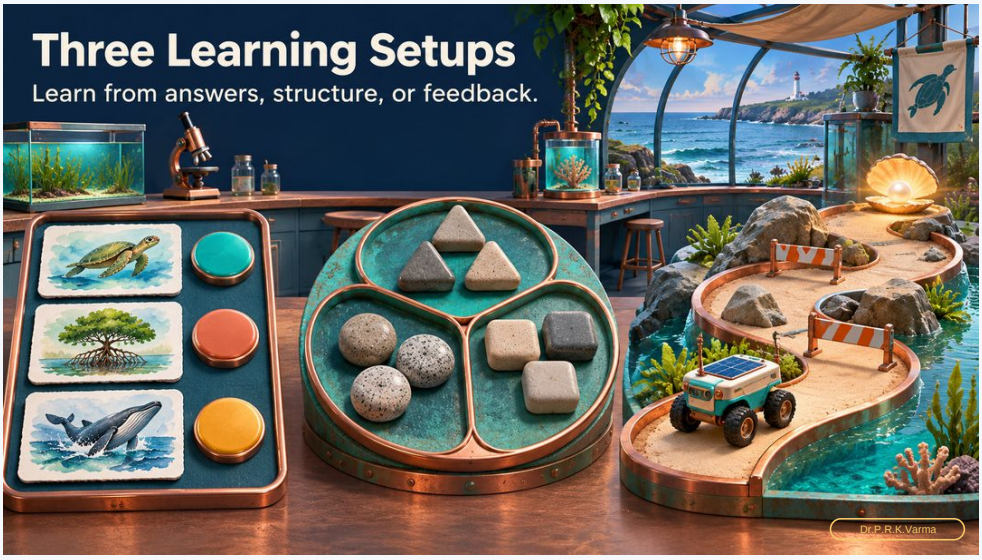
Identify what one example means: a bottle, a site, or a visit, for instance. If several bottles come from the same site and visit, their measurements may be related.



Training Objective and Real Goal

PRESENTER NOTES

A training objective tells the model which predictions to favor. The real goal may include consequences the objective does not capture, so check whether improving the score actually improves the outcome you care about.



Three Learning Setups

PRESENTER NOTES

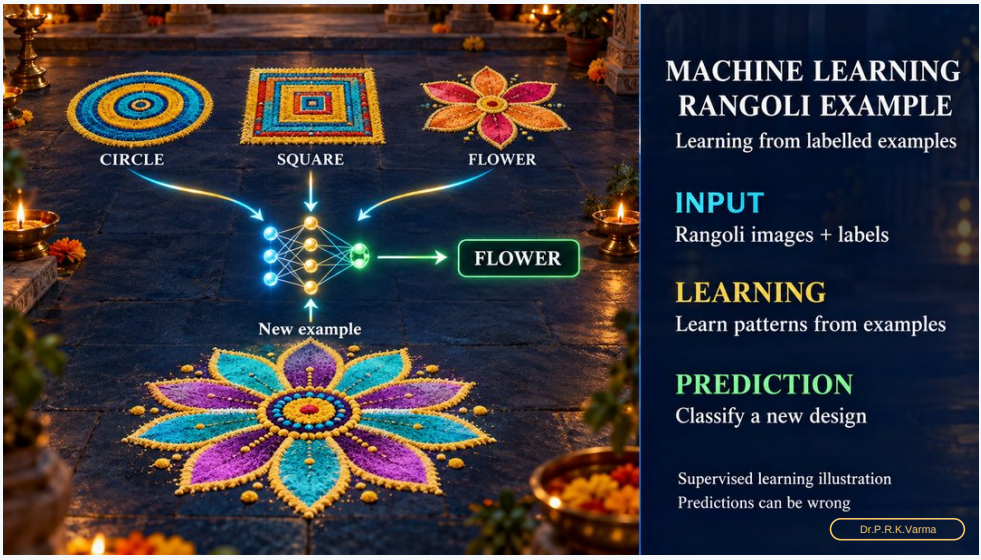
Supervised learning uses target examples, unsupervised learning finds structure without target labels, and reinforcement learning uses feedback from decisions in an environment. The choice depends on the problem and available evidence.



Learning from Labelled Fruit

PRESENTER NOTES

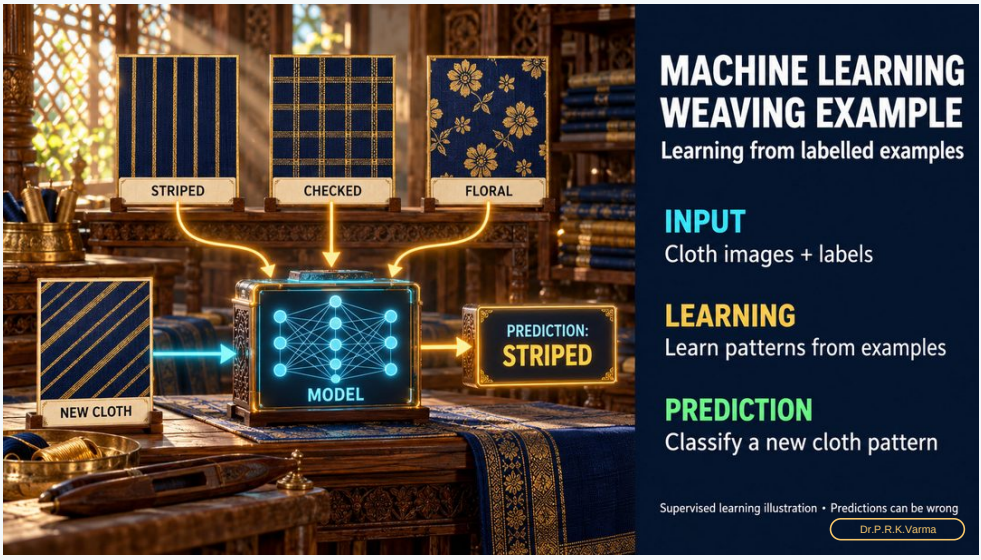
Apple, orange, banana, mango and grapes show how examples can be paired with labels. This is supervised machine learning, one form of ML rather than a definition of every learning method.



Classifying Rangoli Patterns

PRESENTER NOTES

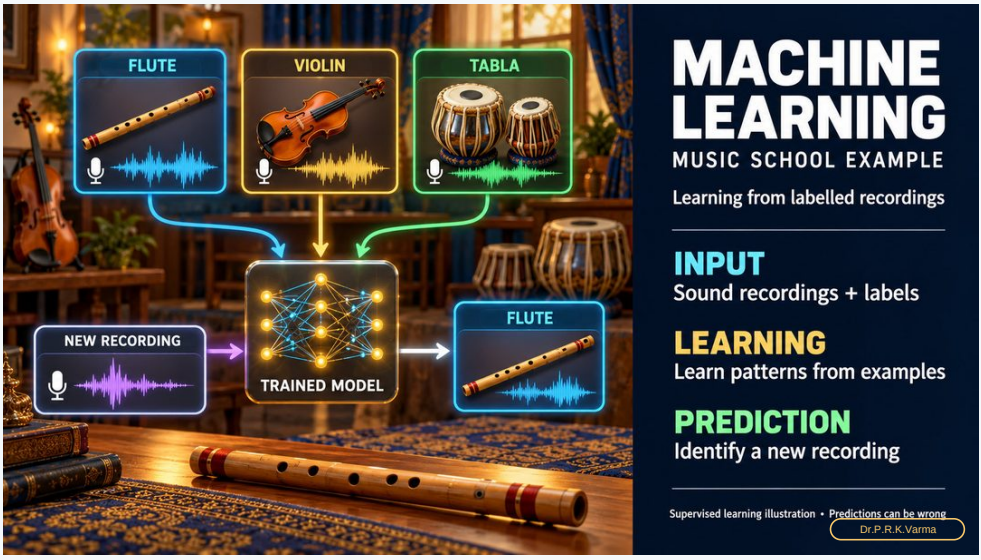
Circle, square and flower labels connect known designs to categories. A learned classifier predicts a category for a new rangoli; a prediction can still be wrong.



Recognising Cloth Patterns

PRESENTER NOTES

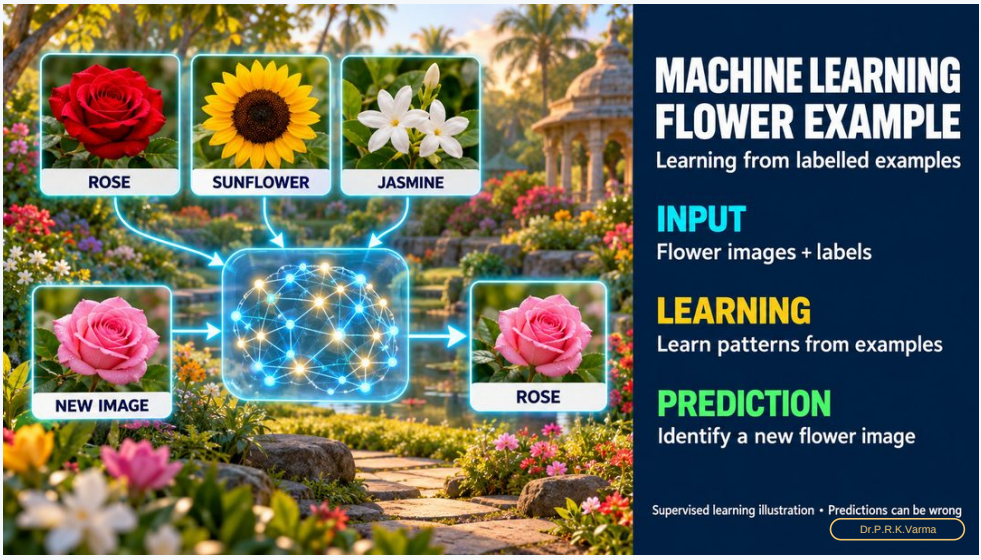
Labelled striped, checked and floral cloth examples teach a classification task. A new cloth image is classified using learned patterns, rather than an explicit hand-written rule for every possible design.



Learning to Identify Instruments

PRESENTER NOTES

Recordings labelled flute, violin and tabla provide training examples. The model then predicts the instrument in a new recording; this illustrates sound classification, not music generation.



Learning Flower Categories

PRESENTER NOTES

Rose, sunflower and jasmine images with labels form a simple training set. The model predicts a label for a new flower image, so varied examples and checks on unseen images matter.



Learning Book Categories

PRESENTER NOTES

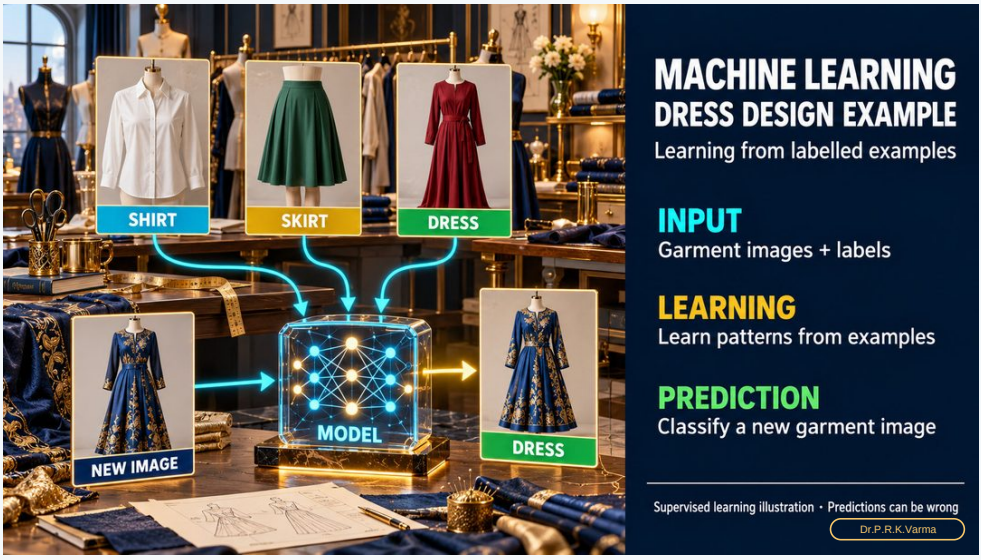
Descriptions labelled science, history and fiction teach a model to associate language patterns with categories. A new description is classified; the output is a category, not a newly written book.



Learning Photograph Categories

PRESENTER NOTES

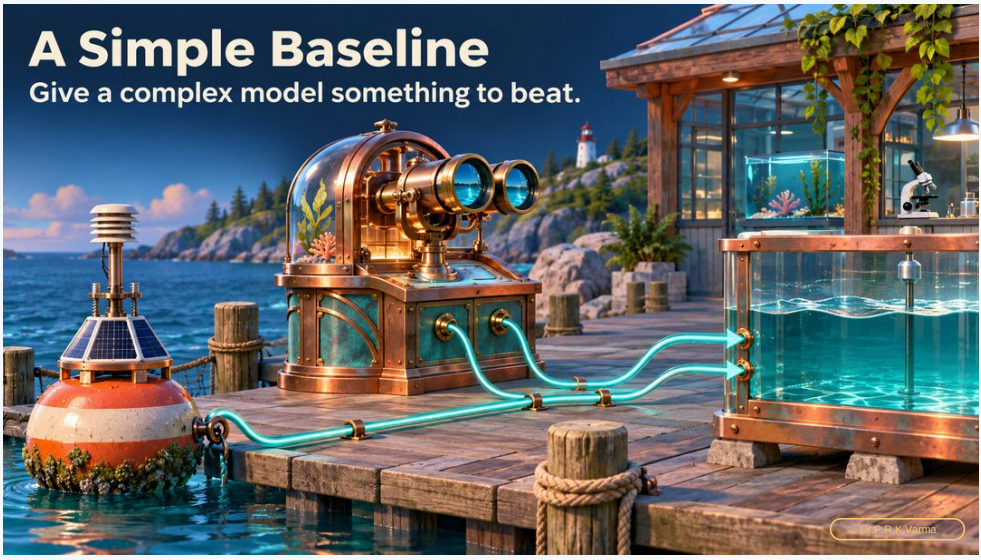
Bird, flower and car photographs illustrate labelled training data. A classifier predicts the category of a new photograph, with reliability depending on whether it learned useful features rather than incidental background clues.



Learning Garment Categories

PRESENTER NOTES

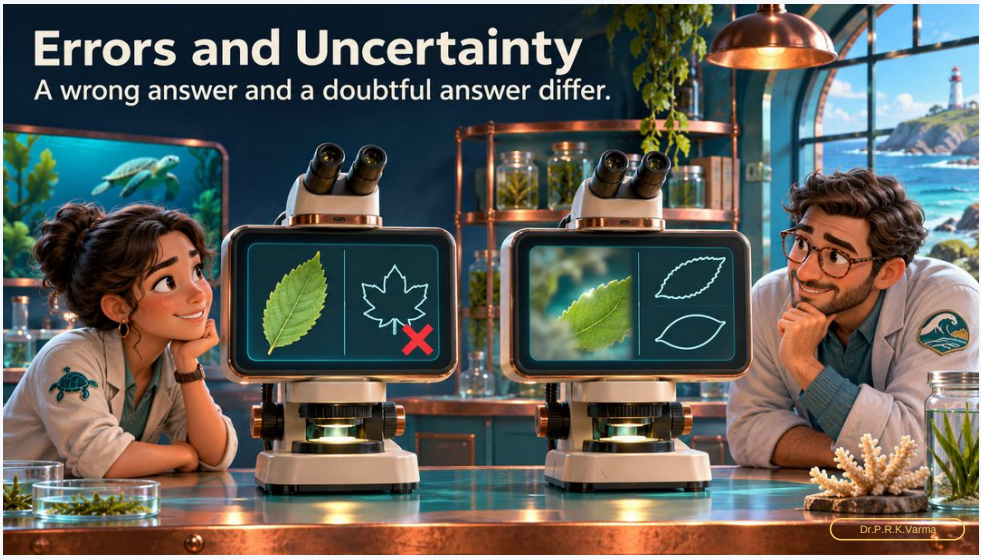
Shirt, skirt and dress photographs provide labelled examples. The model classifies a new garment image; it is not designing a new garment in this task.



A Simple Baseline

PRESENTER NOTES

A baseline might predict the most common outcome, reuse the last measurement, or apply a simple rule. Compare it and the learned model on the same held-out examples.



Errors and Uncertainty

PRESENTER NOTES

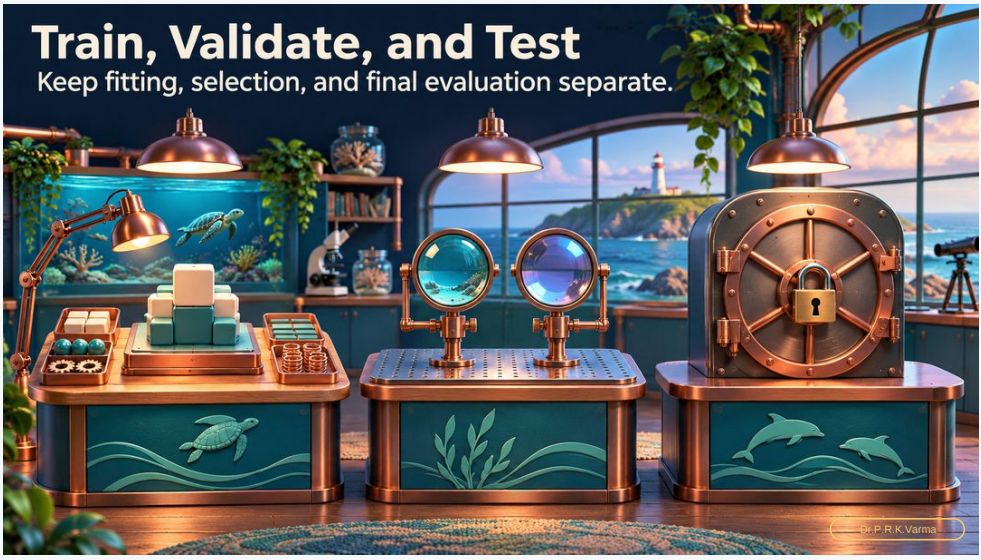
An error is a prediction known to disagree with the reference. Uncertainty is doubt before the reference is known. A model can be wrong with high confidence or right while uncertain.



Errors by Subgroup

PRESENTER NOTES

Report performance for relevant subgroups such as different seasons, sites, or equipment. A strong overall result can mask a weak group. Read small groups carefully because their estimates may be noisy.



Train, Validate, and Test

PRESENTER NOTES

Use training data to fit the model, validation data to tune or select it, and test data for final evaluation. Repeatedly tuning against the test set makes it less useful as an independent check.



Group-Aware Splits

PRESENTER NOTES

If records share a person, device, batch, or site, randomly splitting rows may leak clues across sets. Split by the relevant group when testing performance on unseen groups.



Tea Demand: Learning and Changing Conditions

PRESENTER NOTES

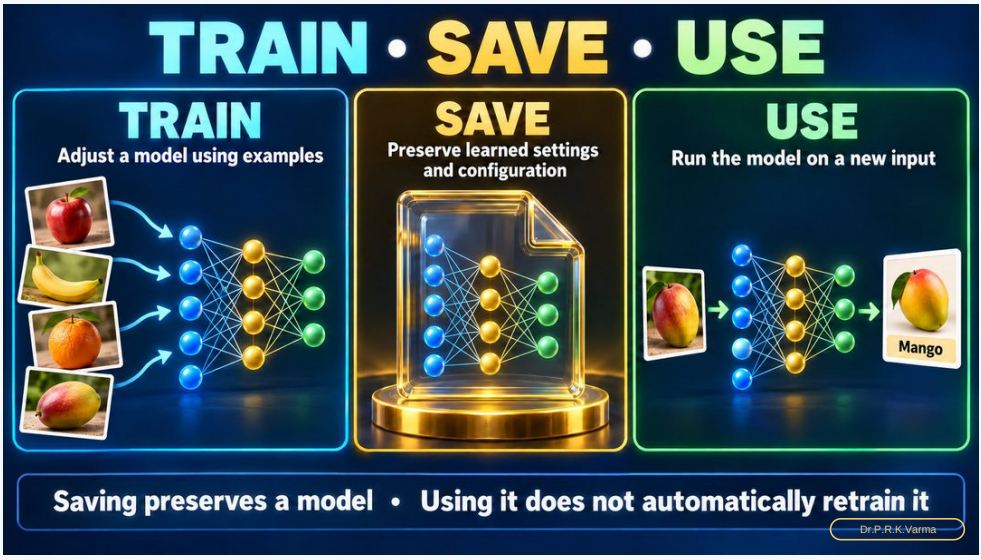
Past tea sales can help estimate future demand. Holidays, closures or other changed conditions may break a previously useful pattern, so a successful model still needs checks in its intended setting.



Generalizing under Change

PRESENTER NOTES

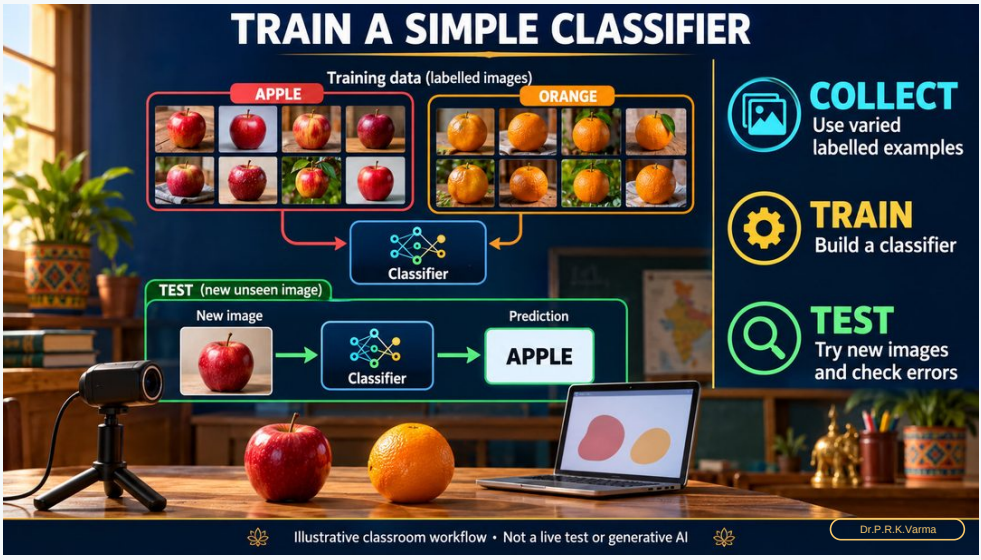
A model can perform well when new inputs resemble training data and struggle when sensors, settings, or populations change. Evaluate in the conditions that matter rather than assuming a held-out score transfers.



Train, Save, Use

PRESENTER NOTES

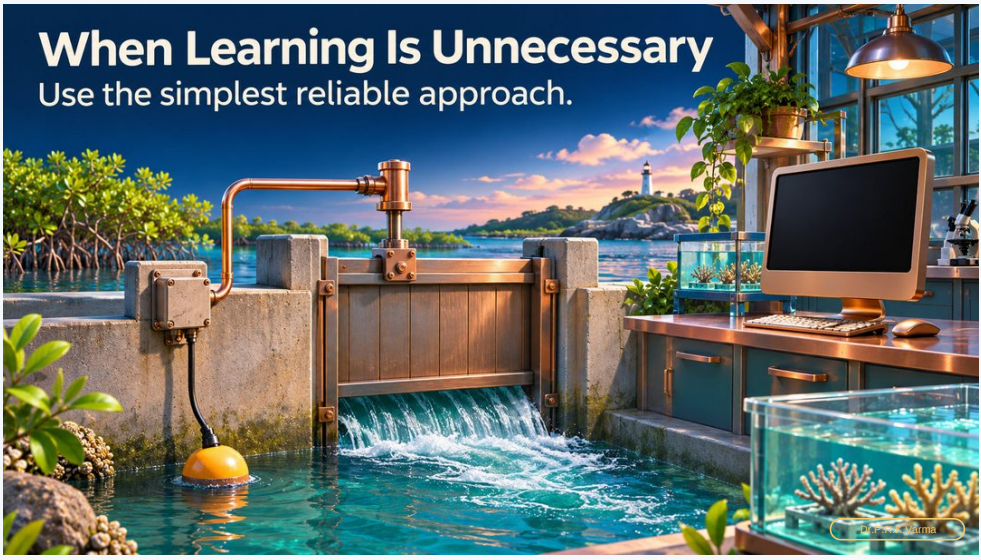
Training adjusts a model using examples. Saving preserves learned settings, and use applies the model to a new input. Saving does not turn an output into a model, and ordinary use does not retrain it.



A Practical Classification Demonstration

PRESENTER NOTES

Train with labelled apple and orange images, then classify a new image. This is a teaching illustration, not the measured result of an experiment that was actually run.



When Learning Is Unnecessary

PRESENTER NOTES

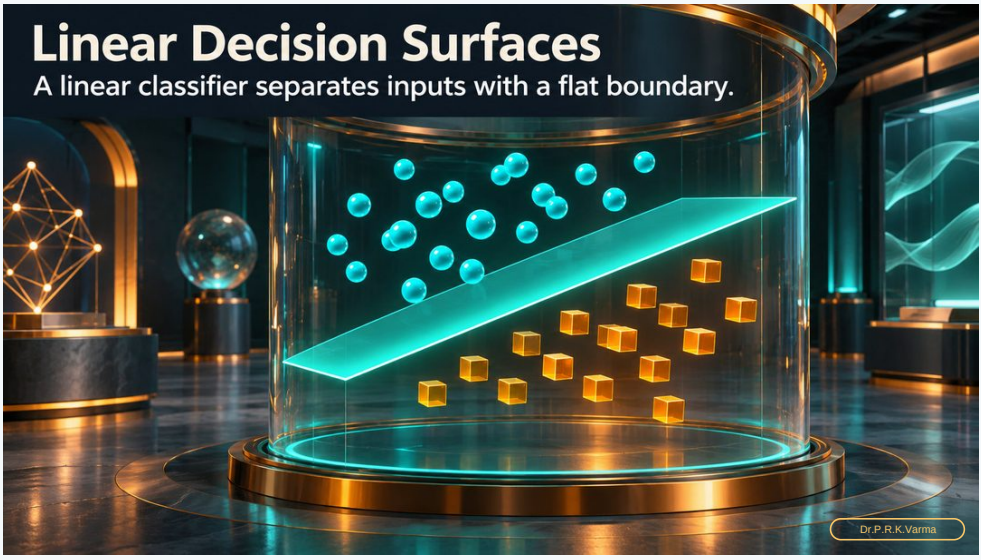
If the requirement is fully specified, stable, and easy to implement, a deterministic rule may be cheaper, easier to verify, and more reliable than a learned model. Reserve learning for patterns that are hard to specify directly.



Five Ms: A Memory Aid

PRESENTER NOTES

Machines, Many Examples, Many Layers, Making Content and Model for Reuse recall five related concepts. They are memory hooks, not five training stages or a chain of nested categories.



Linear Decision Surfaces

PRESENTER NOTES

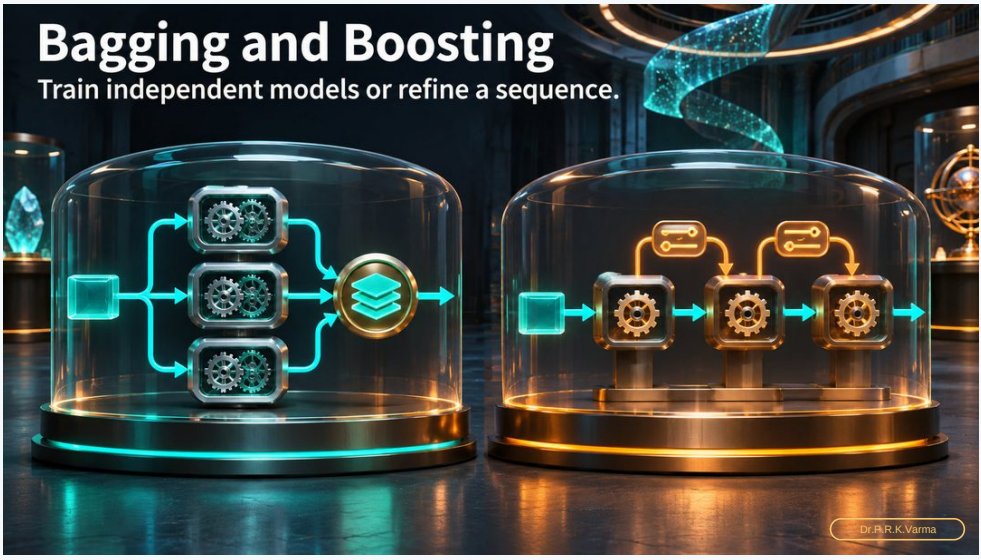
For a binary linear classifier, scores cross the classification threshold at a hyperplane in its feature space. Feature transformations may make the resulting boundary nonlinear in the original inputs.



Decision Trees

PRESENTER NOTES

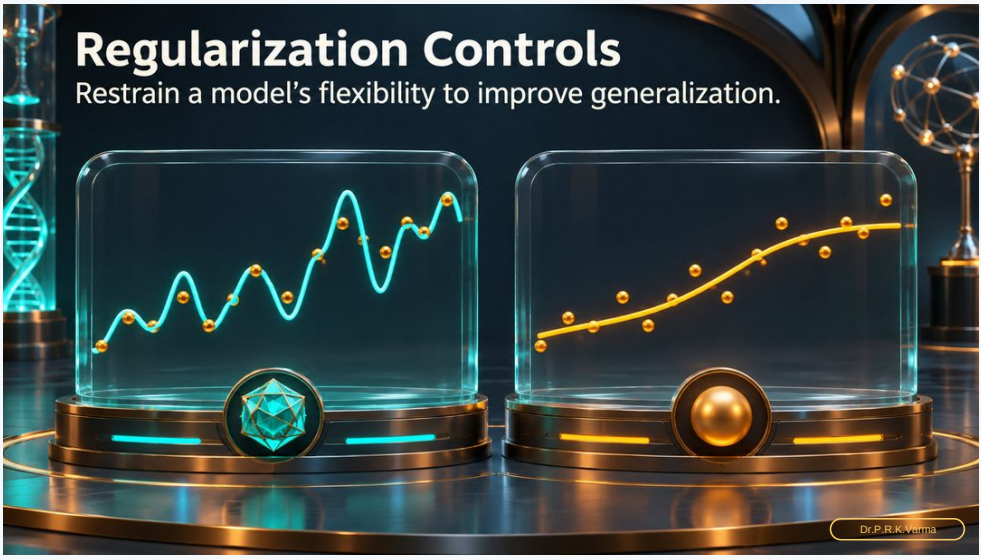
A tree chooses splits that improve its training criterion. At inference, one observation follows a path to a leaf. Deep trees can fit quirks of the training set; control depth or leaf size.



Bagging and Boosting

PRESENTER NOTES

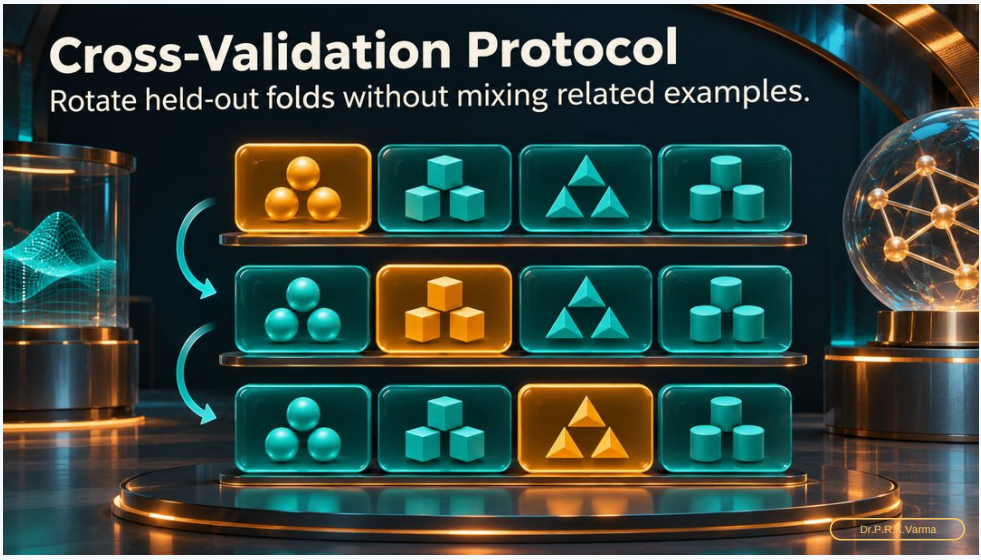
Bagging trains estimators on resampled data and aggregates predictions. Boosting builds a sequence where later estimators improve the current ensemble, often emphasizing hard examples or residuals.



Regularization Controls

PRESENTER NOTES

Regularization may penalize complexity, restrict capacity or stop training early. Tune its strength using validation data; too much restriction can make the model miss useful structure.



Cross-Validation Protocol

PRESENTER NOTES

Train and evaluate separately for each fold, respecting grouping or chronology when relevant. Tune only within the evaluation design and retain an independent test set when making a final selection.



Training-Only Preprocessing

PRESENTER NOTES

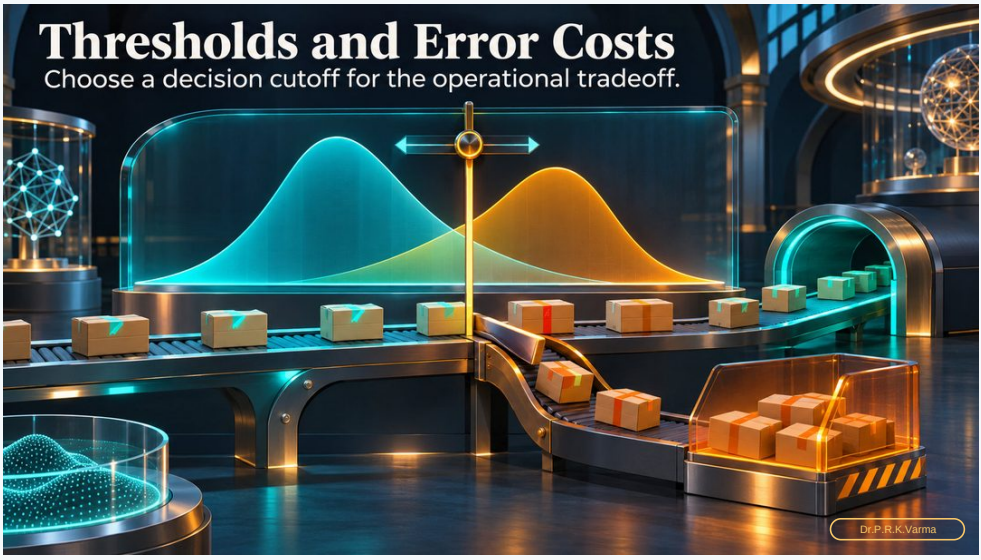
Fit scaling, imputation and other learned transforms only on the training fold, then apply unchanged to validation or test data. Package the preprocessing and estimator together.



Class Imbalance and Weighting

PRESENTER NOTES

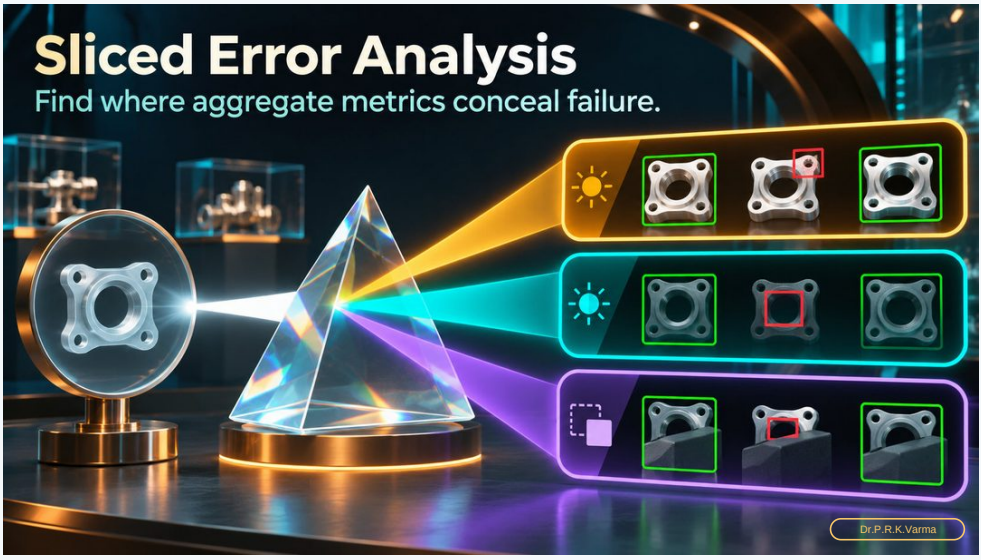
Weighting changes the training objective; resampling changes the data mix seen by the learner. Compare both using relevant precision and recall metrics on a representative validation set.



Thresholds and Error Costs

PRESENTER NOTES

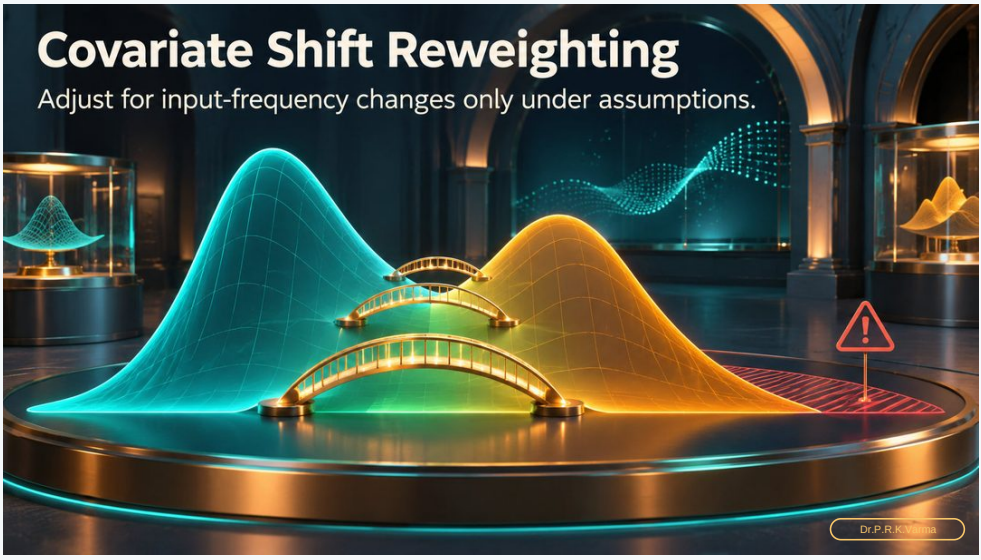
Select a decision threshold on validation examples using false-positive and false-negative costs, capacity or an operating target. Measure the final threshold on independent data.



Sliced Error Analysis

PRESENTER NOTES

Define clinically or operationally meaningful subgroups before looking at results when possible. Report support and uncertainty, inspect representative errors, and verify improvements on independent data.



Covariate Shift Reweighting

PRESENTER NOTES

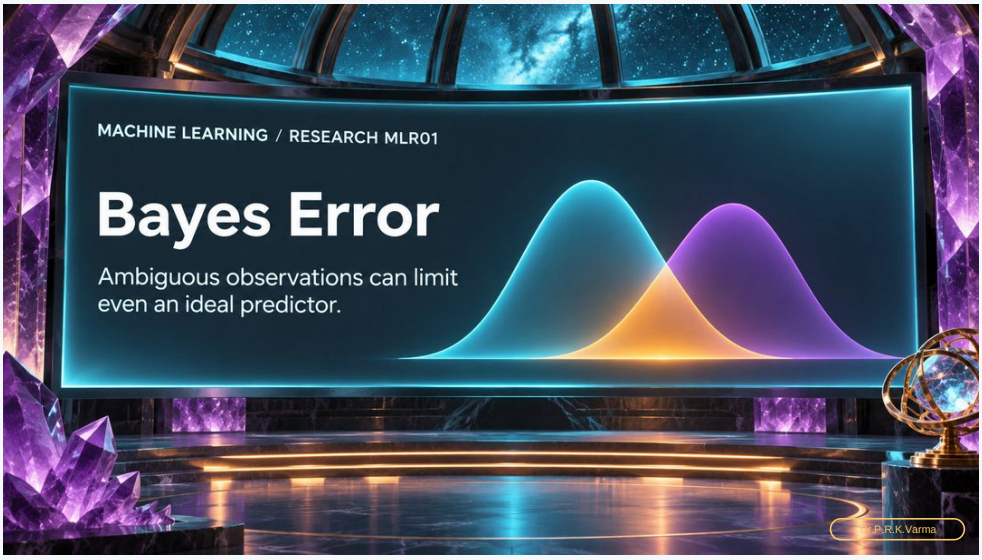
Covariate shift assumes the input distribution changes while the outcome conditional on inputs is stable. Reweighting depends on overlap and reliable weights; unstable or missing overlap limits its usefulness.



Staged Practical Validation

PRESENTER NOTES

Use a held-out offline baseline, monitor predictions in shadow mode before affecting decisions, then consider a bounded pilot with predefined success and rollback criteria.



Bayes Error

PRESENTER NOTES

Bayes error is the lowest achievable average classification error for the assumed data distribution, input features, and loss. Overlapping class possibilities can leave ambiguity even for an ideal decision rule; better inputs may change that limit.

Original research

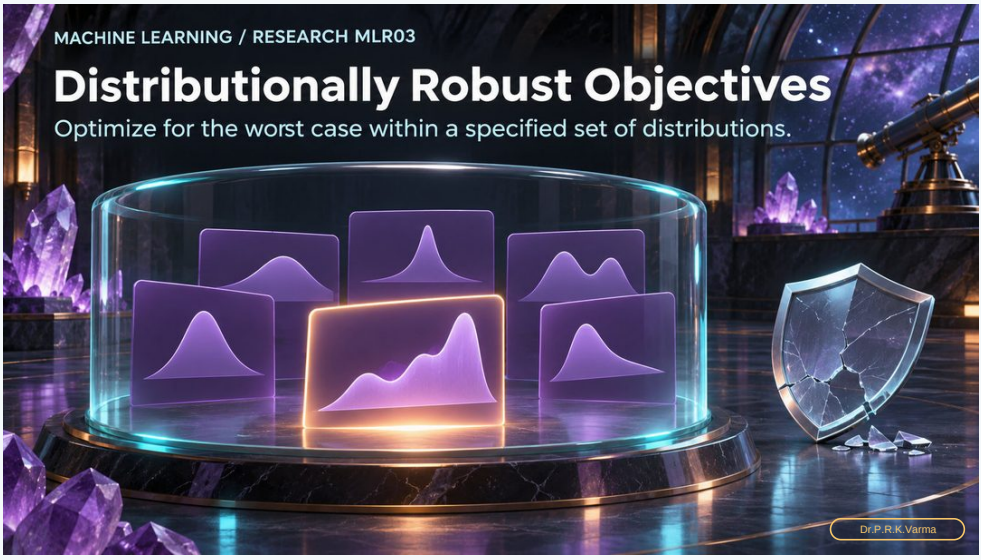


No Free Lunch

PRESENTER NOTES

No-free-lunch results compare learning algorithms under particular averages over possible problems. They explain why useful priors and task structure matter, not why all methods should perform equally on real data.

Original research



Distributionally Robust Objectives

PRESENTER NOTES

Distributionally robust optimization evaluates a predictor across an uncertainty set rather than just one fitted distribution. Specify that set and its scope: protection within it is not universal protection against unknown shifts.

Original research



Selective Labels

PRESENTER NOTES

In the toy factory, a destructive inspection reveals outcomes only for selected parts. Treating uninspected items as successful introduces bias. Gather evidence under a defensible sampling design or clearly state the unobserved-outcome assumptions.

Original research

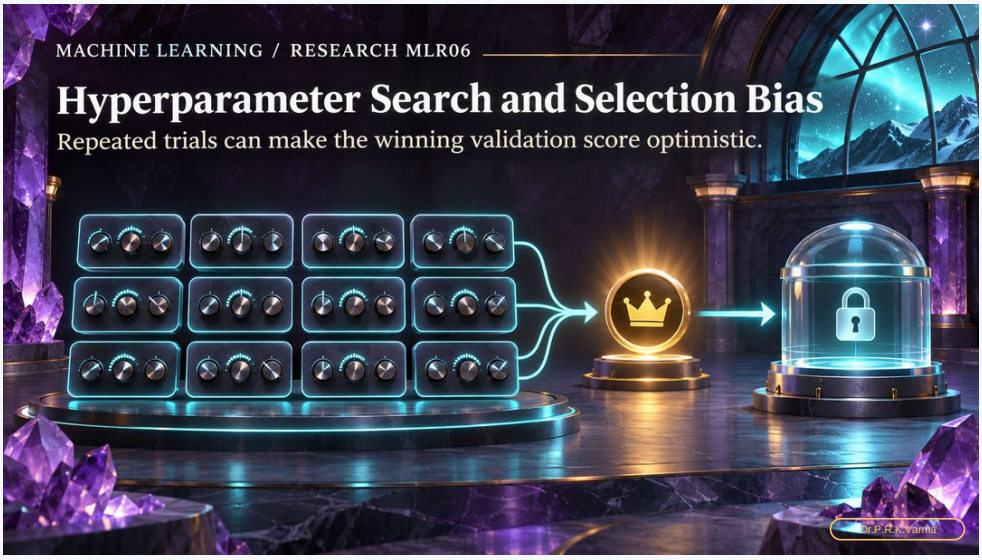


Common Support

PRESENTER NOTES

Common support means the populations being compared both have observations in relevant regions. Limited overlap makes estimates fragile; restrictions can improve comparability but may change the population addressed by the result.

Original research

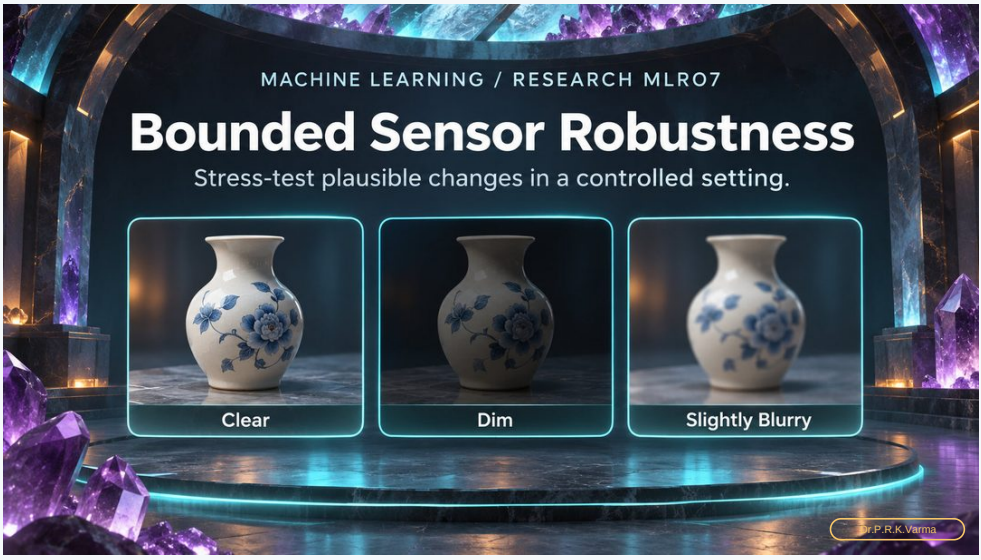


Hyperparameter Search and Selection Bias

PRESENTER NOTES

Selecting the best among many validation trials can exploit random noise in the validation data. Estimate performance in an independent outer evaluation or test set that was not used for tuning.

Original research

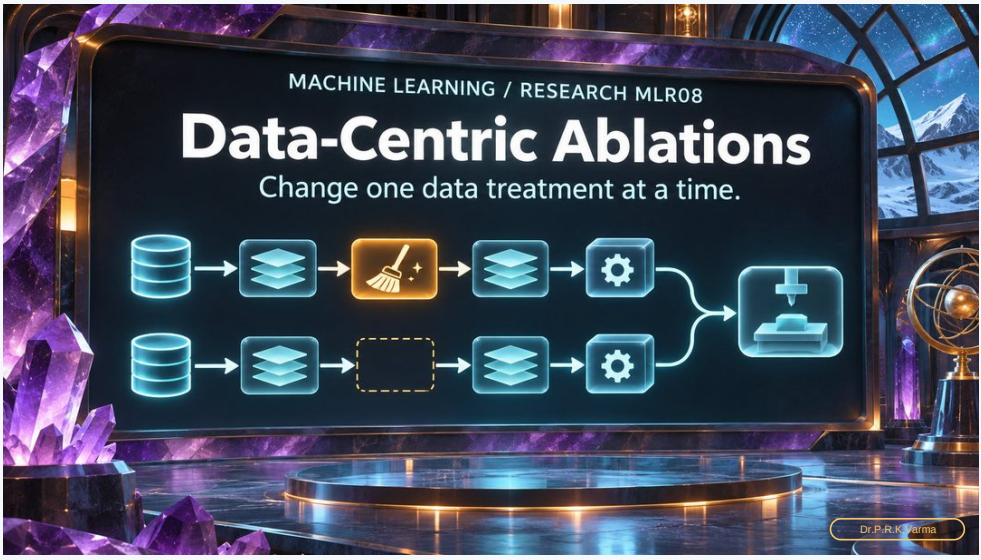


Bounded Sensor Robustness

PRESENTER NOTES

Test plausible changes to a harmless sensor, such as lower illumination or mild blur, at predeclared intensities. Report clean and changed conditions separately; passing these tests does not establish robustness to all changes.

Original research



Data-Centric Ablations

PRESENTER NOTES

Vary one data treatment, such as duplicate removal, and hold model, training budget and evaluation fixed where possible. Repeat the comparison across runs; investigate whether the intervention changed the evaluated population.

Original research

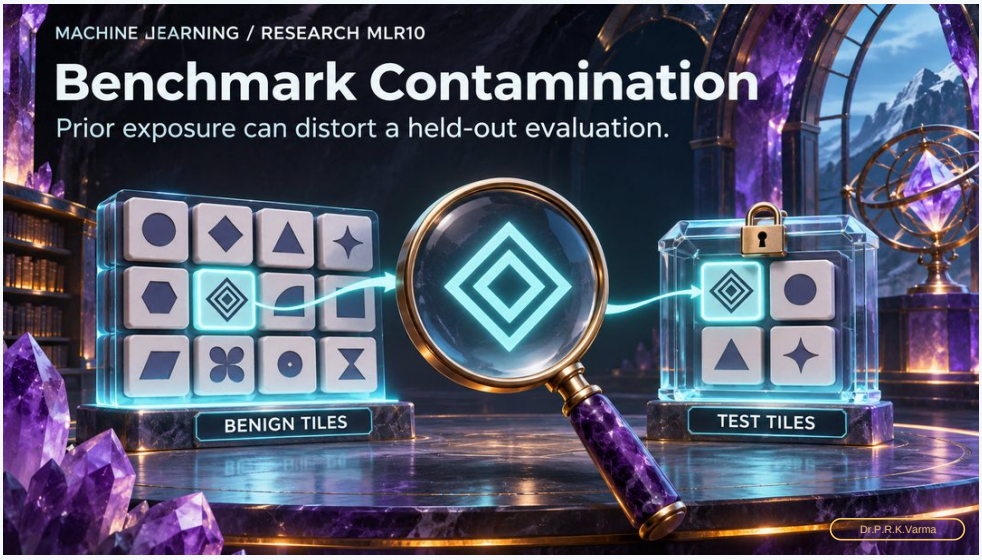


Fairness Metric Trade-offs

PRESENTER NOTES

In abstract groups with different underlying rates, certain calibration and error-balance criteria can conflict unless special conditions hold. State which metric is used and the assumptions behind the comparison.

Original research

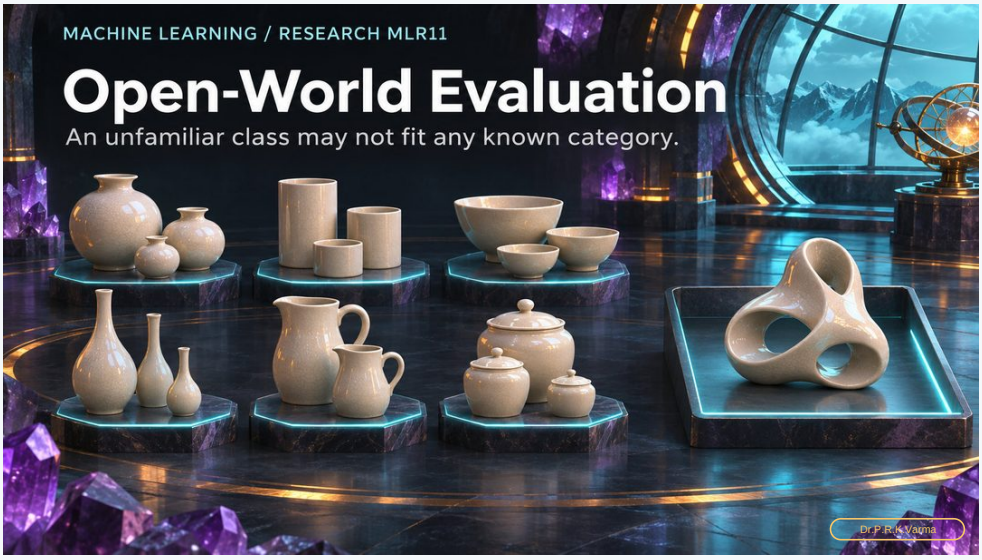


Benchmark Contamination

PRESENTER NOTES

A benchmark is less informative if near-duplicate content was previously seen. Check overlap and state the detection limitations; compare performance on a separately collected set when overlap cannot be ruled out.

Original research



Open-World Evaluation

PRESENTER NOTES

Open-set evaluation includes classes absent from model training. Assess recognition of known classes and rejection or flagging of novel inputs, instead of forcing every example into the training vocabulary.

Original research