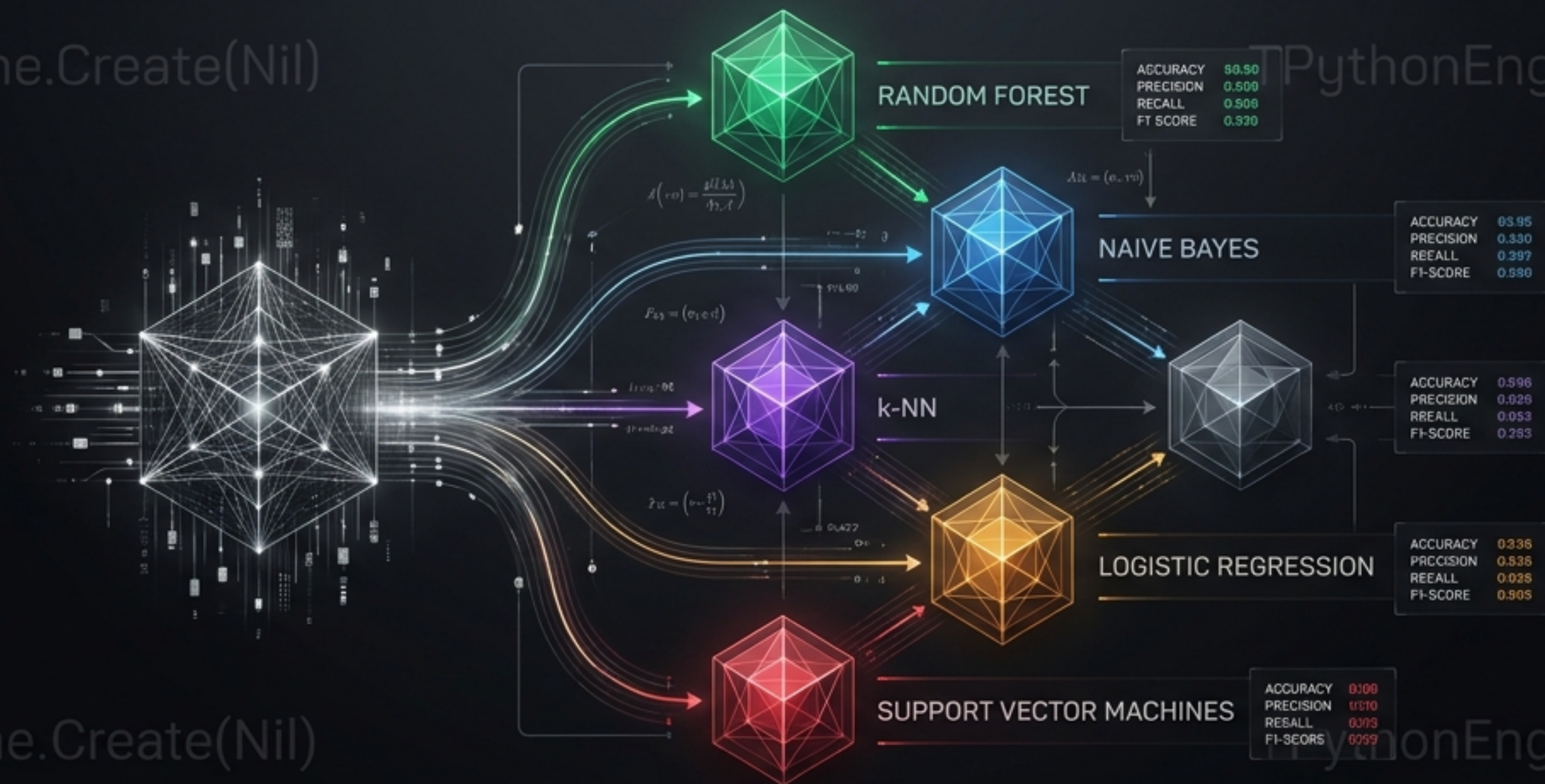


ROBUST MACHINE LEARNING IN MAXBOX

MULTI-CLASSIFIER BENCHMARKING & ADVANCED CROSS-VALIDATION STRATEGIES

TPythonEngine.Create(Nil)

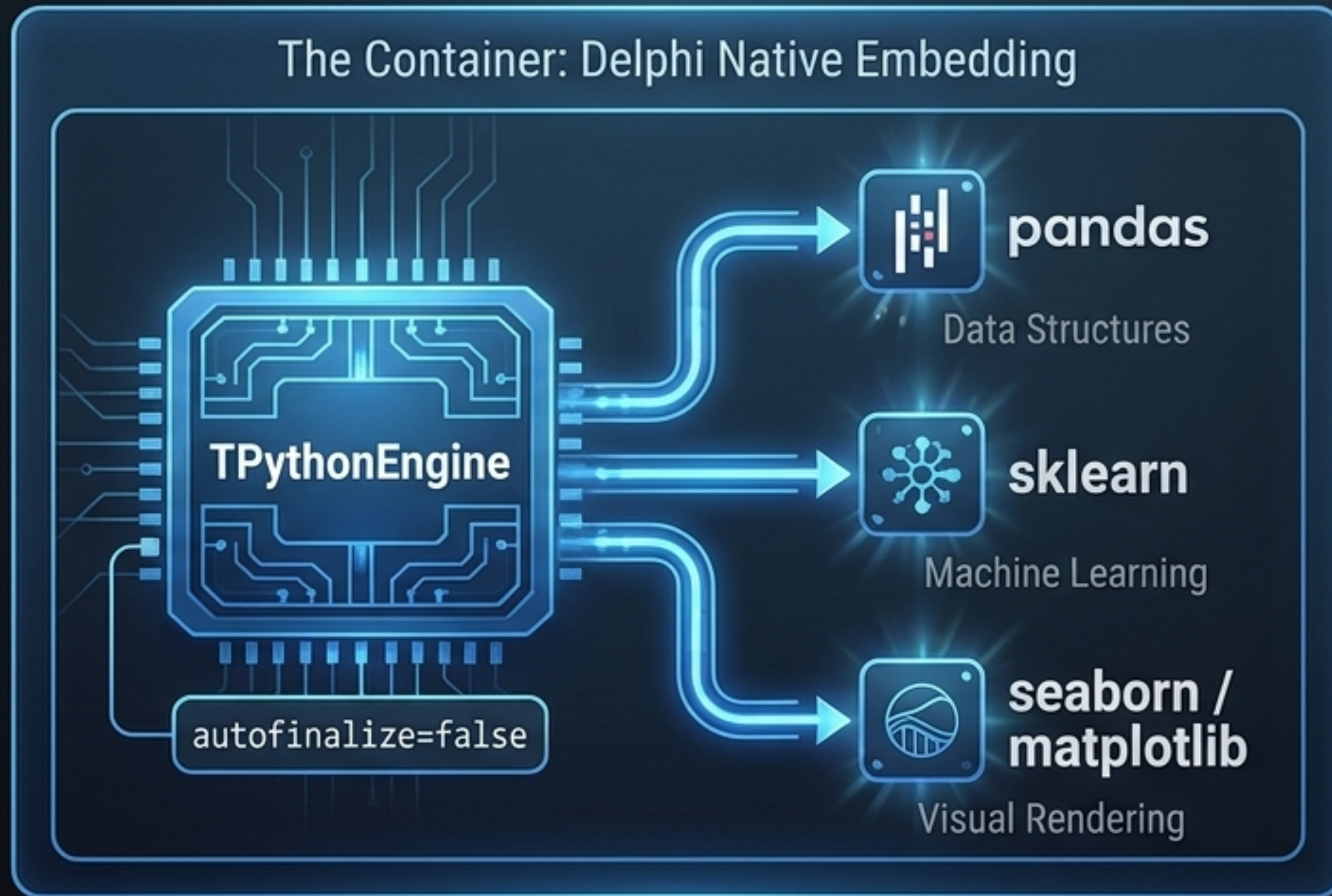


TPythonEngine.Create(Nil)

DATA			
	ACCURACY	PRECISION	F1 SCORE
1	0.95	0.90	0.92
2	0.89	0.83	0.86
3	0.89	0.85	0.87
4	0.80	0.80	0.80
5	0.80	0.80	0.80
6	0.80	0.80	0.80
7	0.80	0.80	0.80
8	0.80	0.80	0.80
9	0.80	0.80	0.80
10	0.80	0.80	0.80



The maXbox Orchestration Layer



```
=====
maXbox Machine Learning Execution
Log
=====
Dataset Selected : Iris Flower
Dataset (150 samples)
Algorithm       : 1. Random Forest
Classifier
Hyperparameter 1 : 100.00000000000000
Hyperparameter 2 : 5.000000000000000

---> Running Random Forest
Classifier (Ensemble Trees)
Building 100.00000000000000 decision
trees with Gini criterion...
Dataset Split: 50% Train (75
samples), 50% Test (75 samples)
Classification Accuracy: 94.67%
Precision (Weighted): 0.948 |
```

Seamless integration of Python's premier data science libraries directly into compiled maXbox desktop applications without heavy overhead.

Scaling the Complexity: From Baseline to Benchmark

The Baseline (Iris Dataset)



Dimensionality:

4 Features

Volume:

150 Samples

Strategy:

Basic 50/50
train_test_split

The Challenge (Wine Dataset)



Dimensionality:

13 Features (Alcohol,
Malic Acid, etc.)

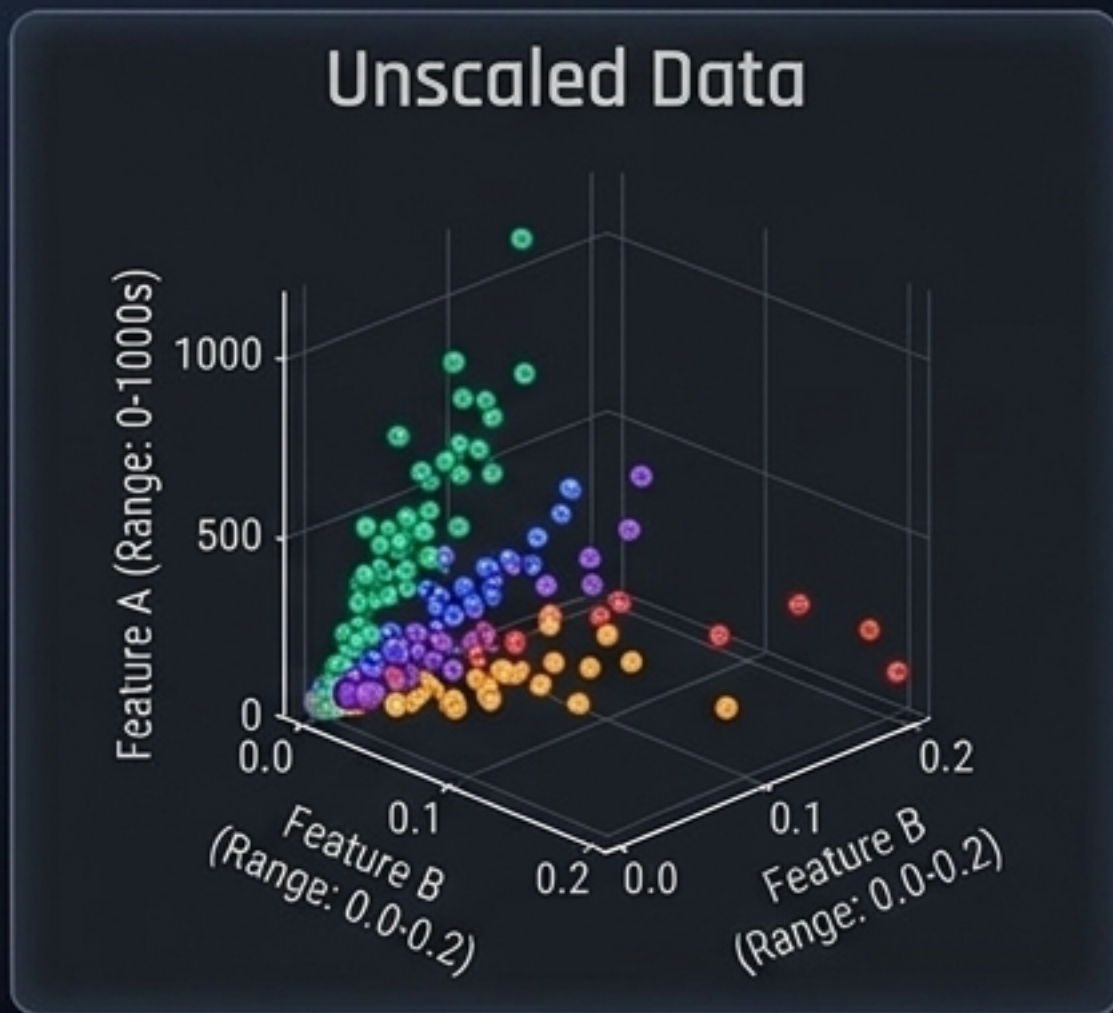
Volume:

178 Samples (3 cultivars)

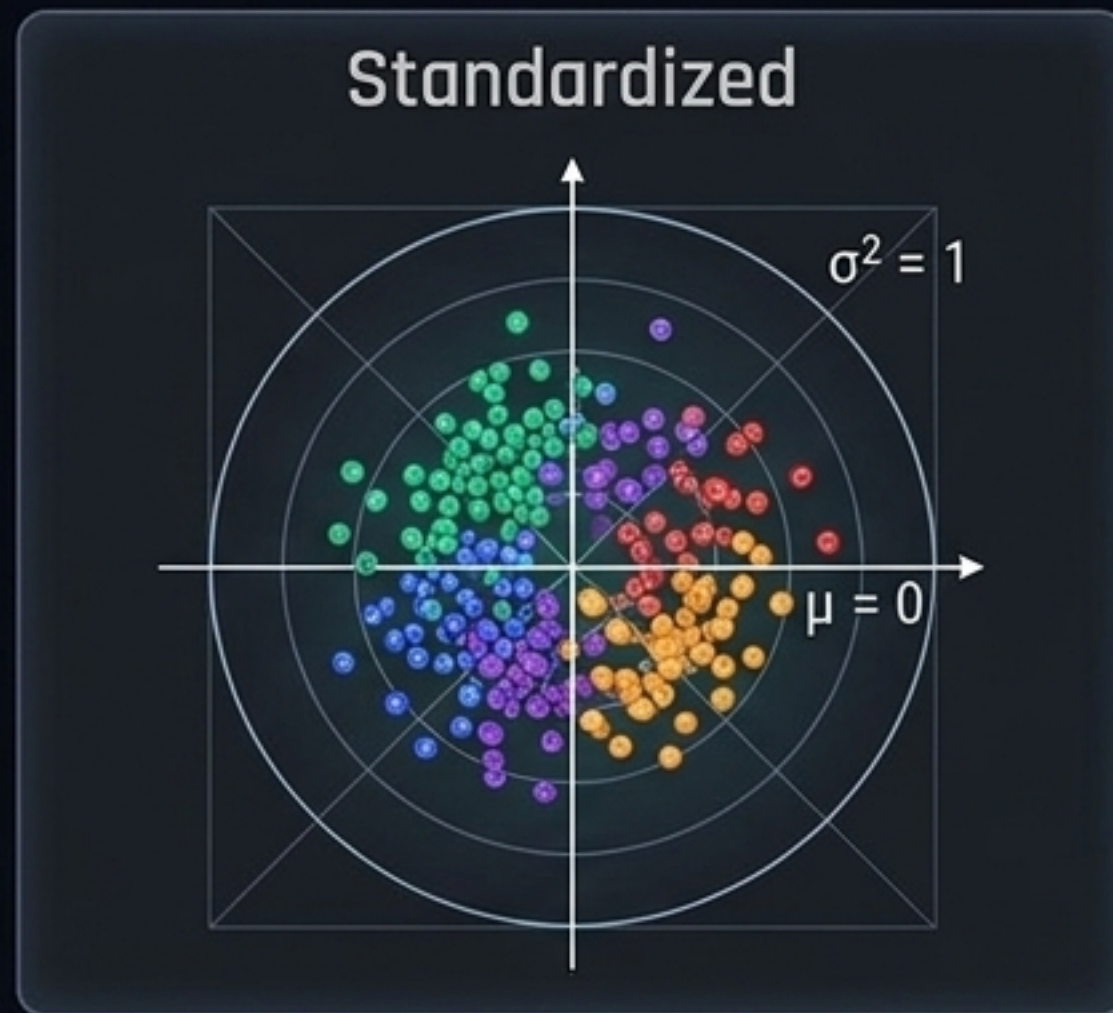
Strategy Required:

Advanced Preprocessing
& Validation

Calibrating the Arena: Feature Scaling with StandardScaler



StandardScaler



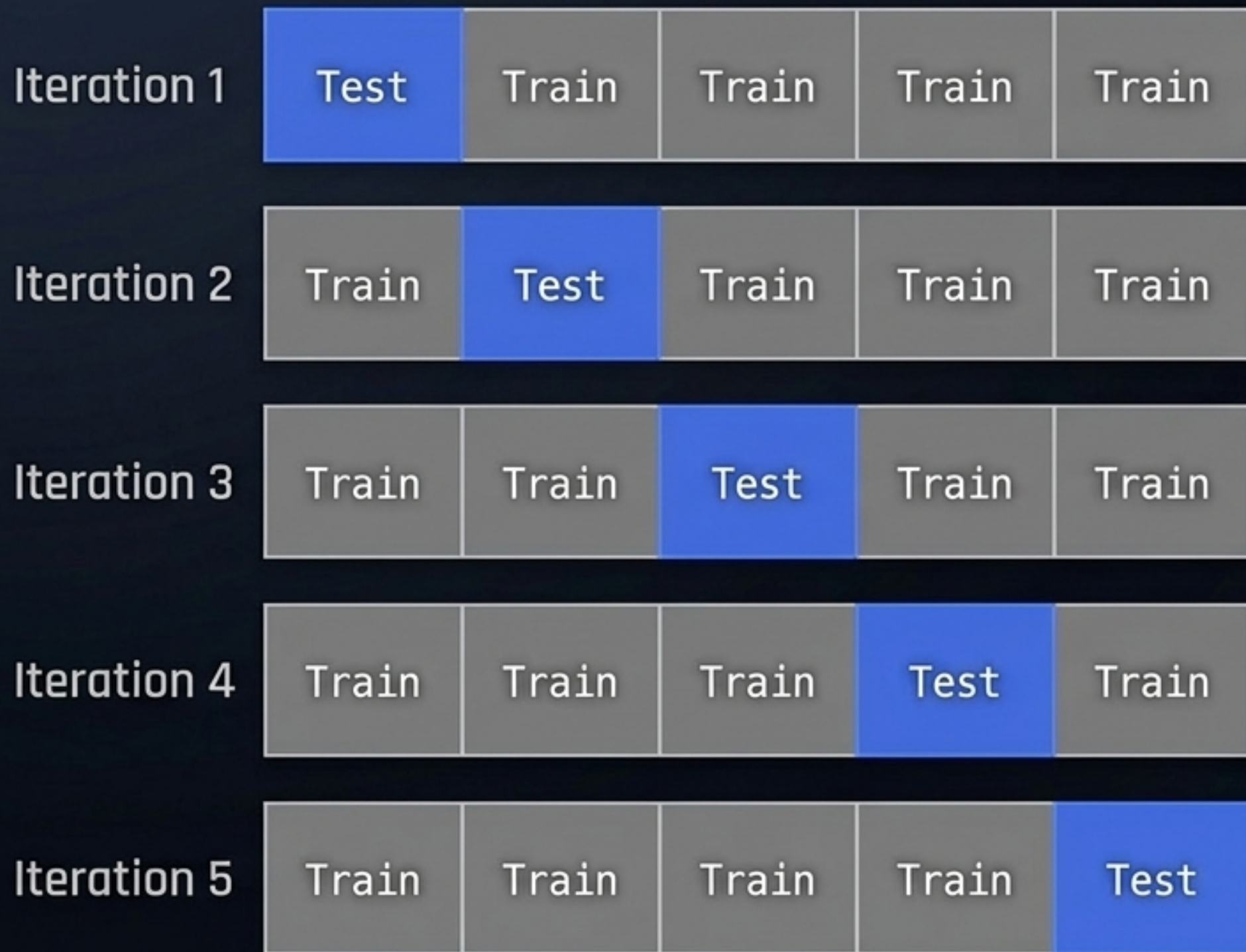
The Problem: Features with massive numerical ranges overpower Euclidean distance calculations.

When features are on different scales, those with larger values dominate the distance computations, leading to biased model learning, particularly in distance-based algorithms like k-NN and SVM.

The Solution: Standardizing 13 features ensures fair evaluation for distance-based models like SVM and k-NN.

StandardScaler transforms data to have a mean of 0 and a standard deviation of 1. This normalization allows all 13 Wine Dataset features to contribute equally to the model's performance.

The Validation Engine: 5-Fold Stratified K-Fold



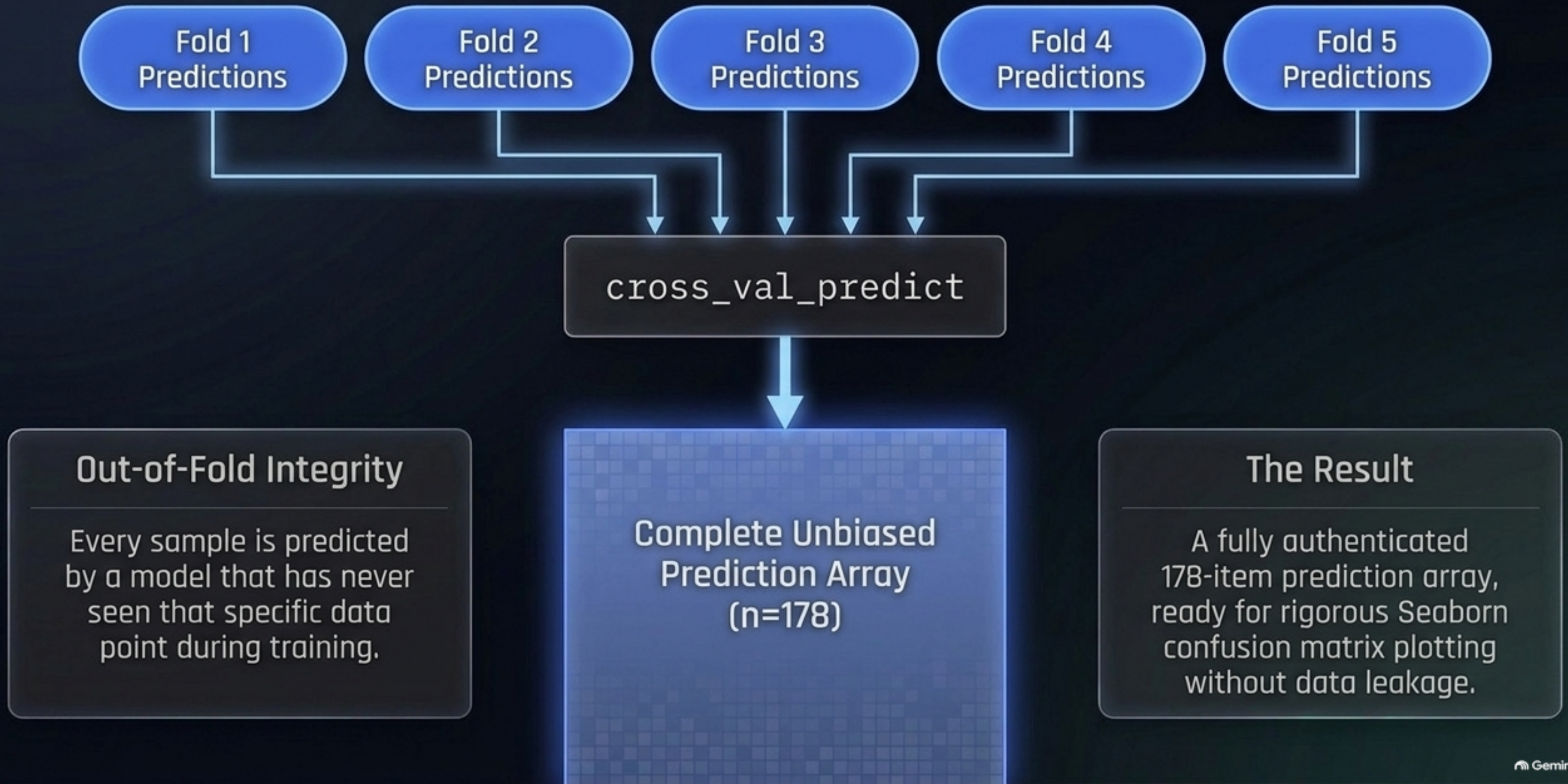
Beyond the 50/50 Split

Models are trained and evaluated 5 separate times, ensuring comprehensive coverage.

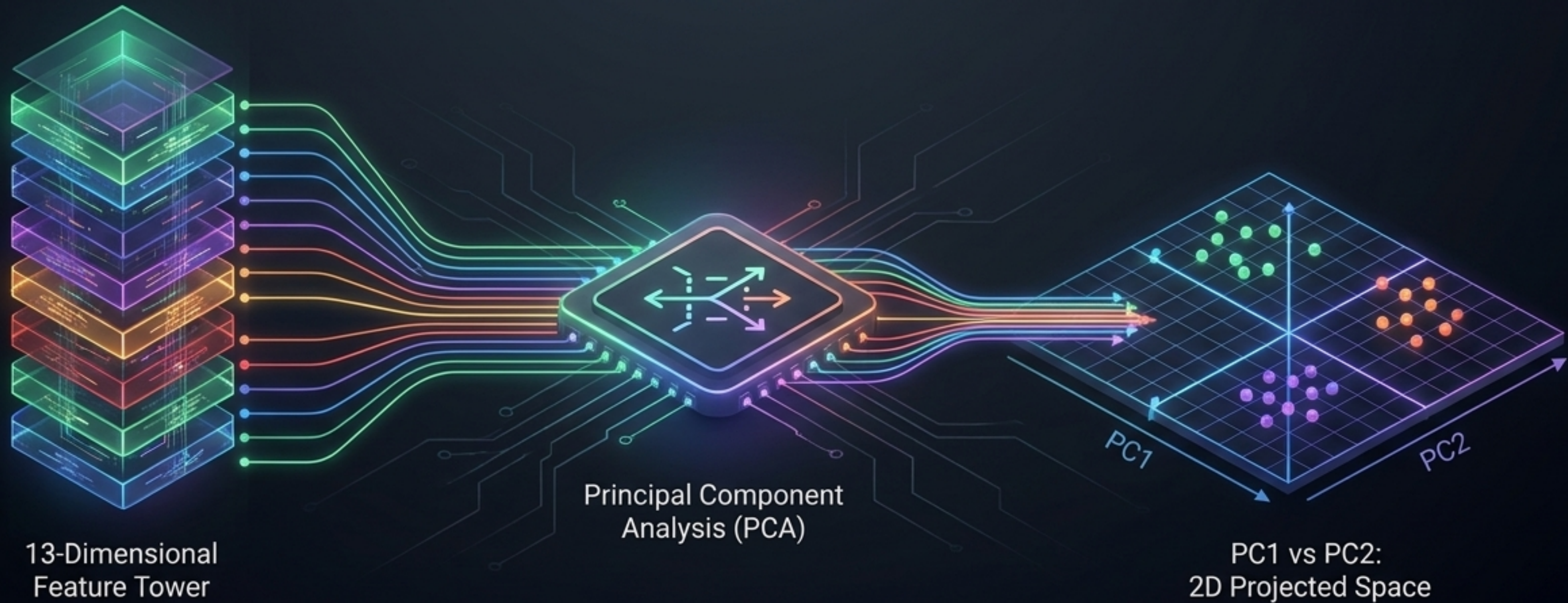
Stratification Guarantee

Preserves the exact original ratio of the 3 Wine cultivars within every single fold. No fold is statistically compromised.

Assembling the Truth: Out-of-Fold Predictions



Dimensionality Collapse: The 2D PCA Canvas



Strategic Distinction:

Models are rigorously trained and cross-validated on the full 13-dimensional dataset. PCA is deployed solely as a lens to draw clear, interpretable 2D decision boundary maps.

The Competitors: Multi-Classifier Diagnostic

Random Forest

Ensemble

Builds multiple decision trees and merges them for stability and variance reduction.

SVM

Margin-based

Finds the optimal hyperplane that maximizes the gap between different classes.

k-NN

Instance-based

Classifies based on the most frequent label among the closest training examples.

Naive Bayes

Probabilistic

Applies Bayes' theorem with the assumption of conditional independence between features.

Logistic Regression

Linear

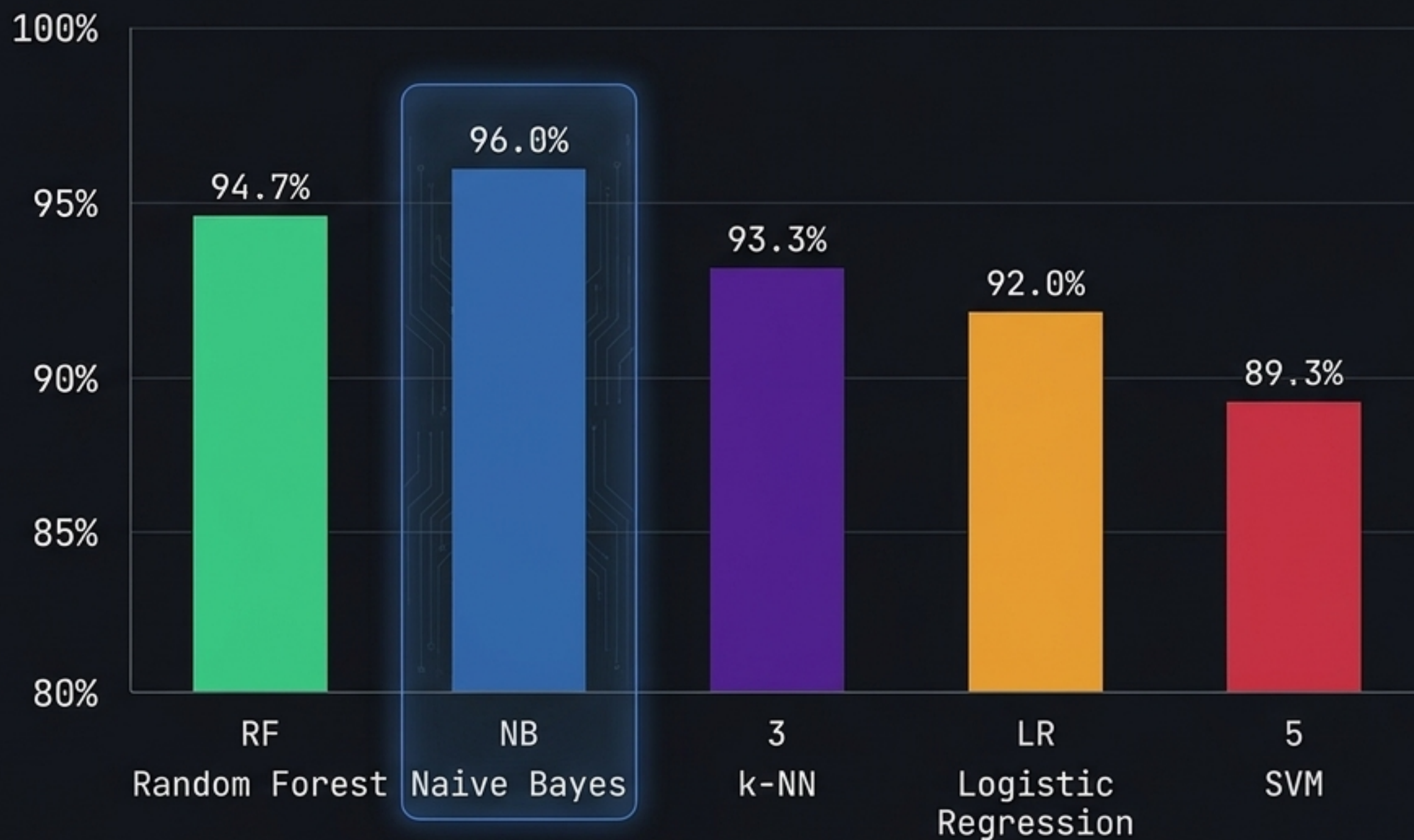
Uses a logistic function to model the probability of a discrete class.

Mapping the Logic: 2D Decision Boundaries



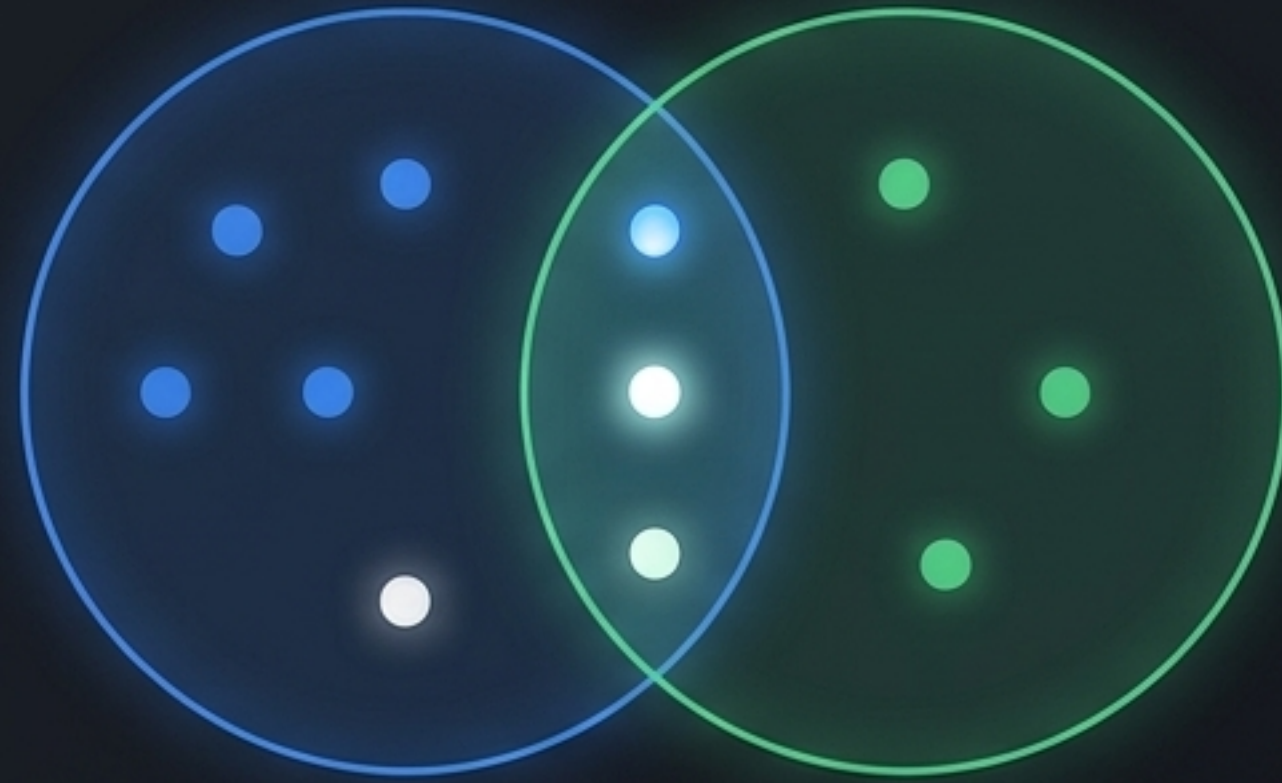
Observe the aggressive, linear cuts of Logistic Regression contrasted against the highly localized, non-linear pockets generated by k-Nearest Neighbors.

The Benchmarking Results: Predictive Accuracy



All models evaluated via 5-Fold Stratified CV on the fully scaled 13-feature Wine dataset.

Diagnostic Realities: Precision vs. Recall



Recall (Catching the target)

The Human Analogy: If you don't recognize a person, it is bad recall.

Wine Dataset: Failing to identify a true Cultivar 2 wine.

Precision (Purity of the catch)

The Human Analogy: If you recognize a person for another (confuse them), it is bad precision.

Wine Dataset: Classifying a wine as Cultivar 1 when it is actually Cultivar 3.

The Matrix Mosaic: Out-of-Fold Error Mapping

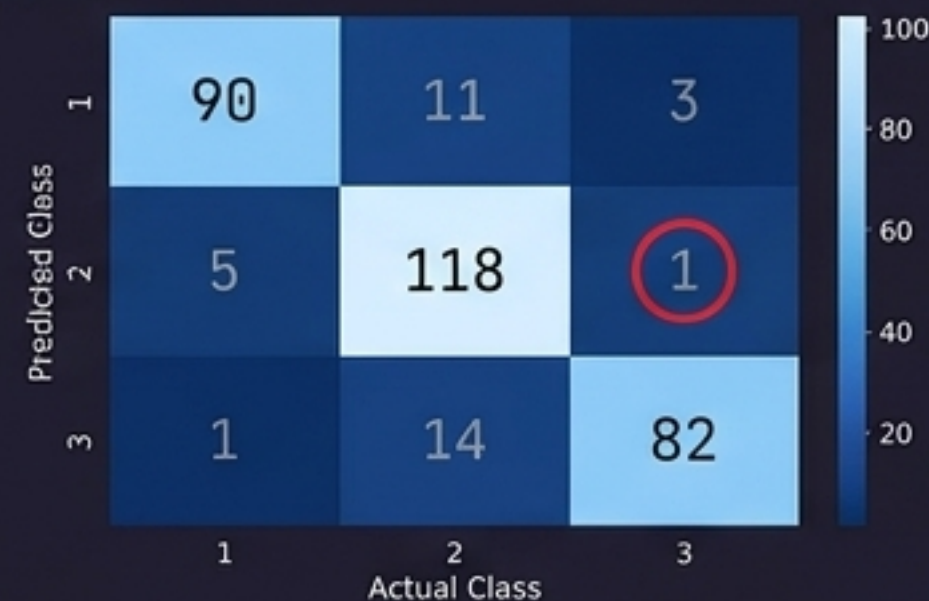
Random Forest



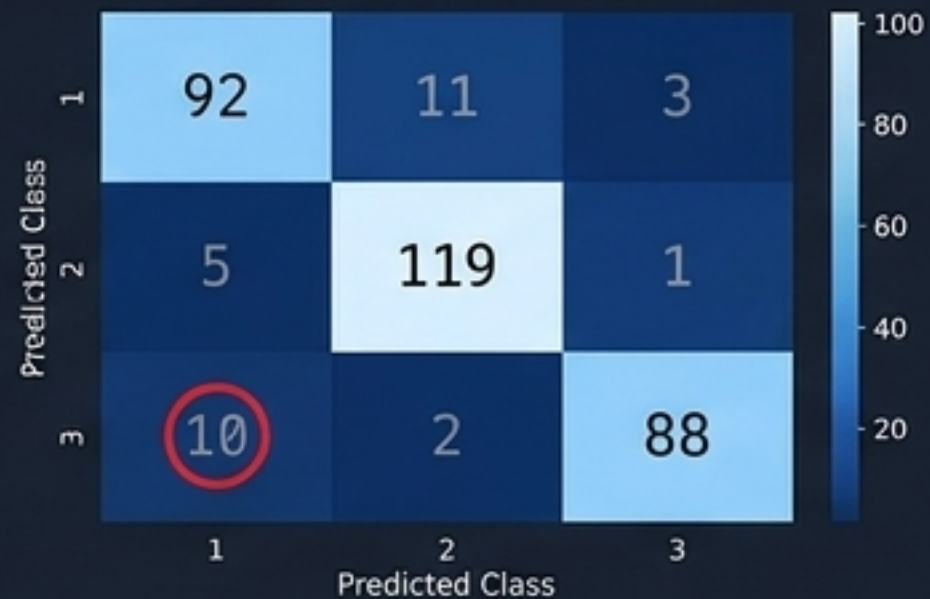
SVM



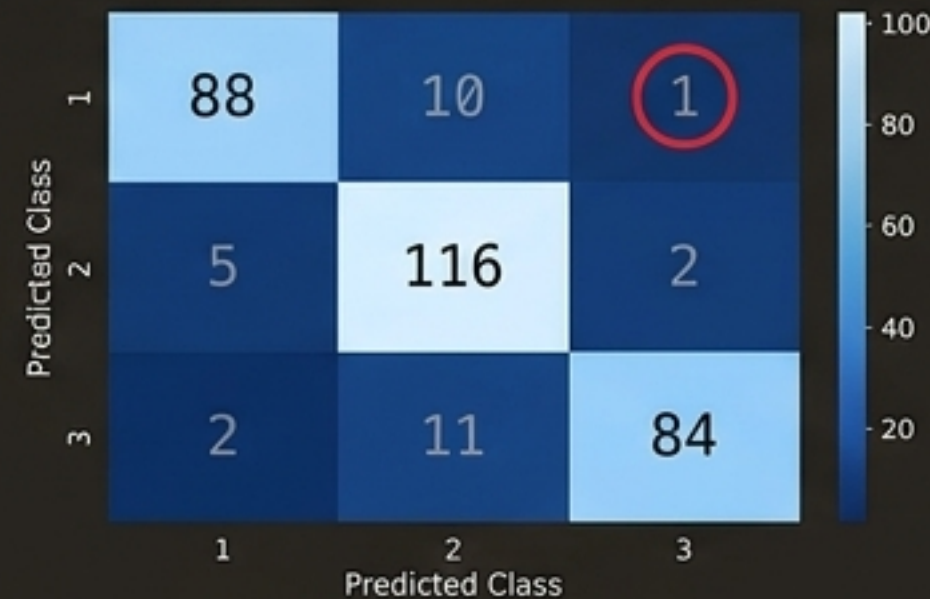
k-NN



Naive Bayes



Logistic Regression



A unified view of model blind spots powered by `seaborn.heatmap` and `cross_val_predict`.

Executive Synthesis: The maXbox Pipeline Verdict

The Verdict

Demonstrated the highest resilience and precision across 13 dimensions.



Data Prep

StandardScaler
algorithmic alignment



Validation

Stratified 5-Fold CV
out-of-fold predictions



Architecture

Native Python4Delphi
execution

The Trust Factor: The result isn't a fluke of a lucky 50/50 split. It is the product of robust, scaled cross-validation executed seamlessly within the maXbox environment.