DISSERTATION RESEARCH

Detection of Population Drift in AI-Powered Medical Devices

A Comparative Analysis of Unsupervised Drift Detection Methods

Presented by : Malik Adeel Anjum

Population Drift

Machine learning models assume the future will look like the past – but it doesn't.

Silent Failure Risk

- Regular software fails, they generate errors and alarms
- When AI model fails, there are no errors just " wrong guess with high confidence"

Patient Safety & Misdiagnosis

- Since model failure is silent, misdiagnoses go unnoticed
- Detrimental to patient safety

MHRA Regulatory Requirements

- MHRA mandates continuous monitoring and robust drift detection for all AI as a Medical Device (AIaMD)

Three Types of Population Drift

Understanding how clinical data evolves over time

Data Drift

Covariate Shift

Changes in input feature distributions over time

Example:

Ageing population, changes in population demographics

Concept Drift

Rules Change

Relationship between features and target evolves

Example:

New biomarkers and revised medical definitions

Prevalence Drift

Disease Frequency

Changes in class distribution over time

Example:

Disease frequency shifts due to epidemics or pandemics

The Research Gap

Critical limitations in current drift detection methods

1 Supervised Methods Inadequate

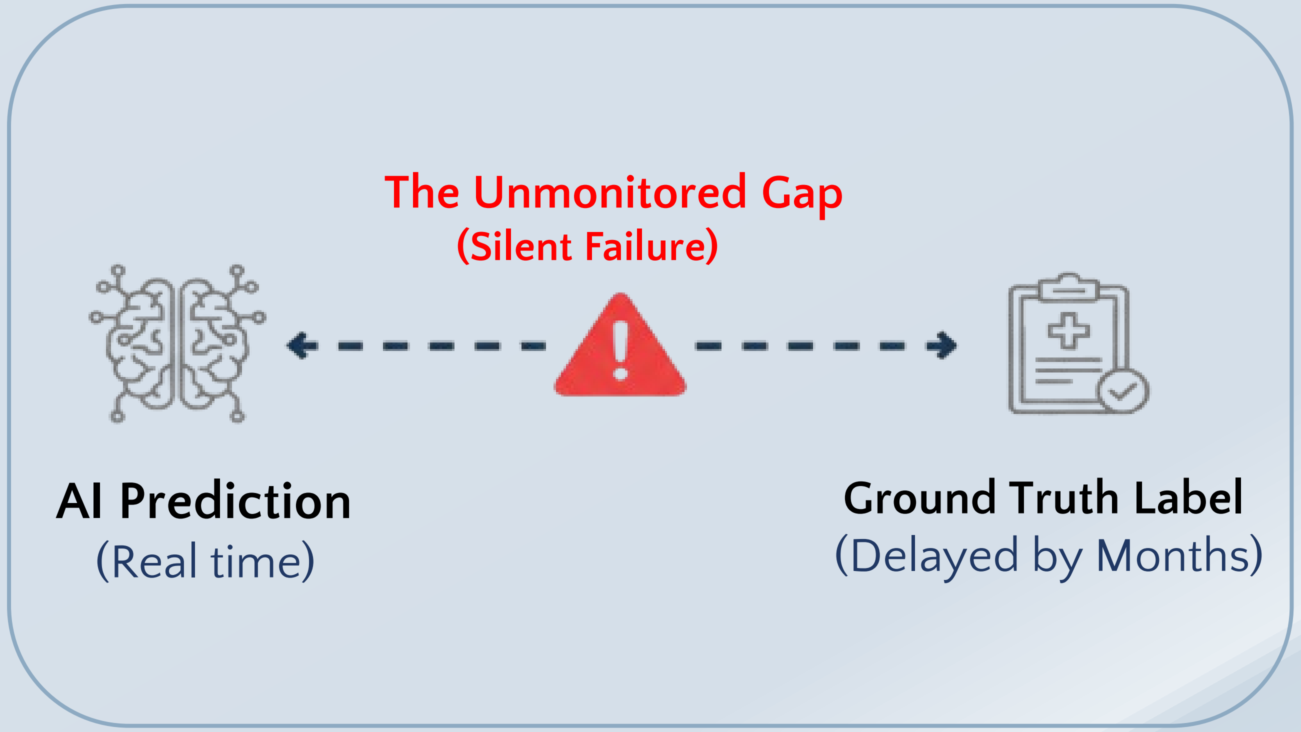
- Current methods (DDM, ADWIN) require immediate labels
- Labels arrive months later
- Real-time monitoring is impossible

2 Limited Direct Comparisons

- Existing research tests single algorithms in isolation
- Missing: Head-to-head benchmarks (Partitioning vs. Distance)

3 Detector-Drift Type Mismatch

- Unclear which detector suits which drift type (Gradual vs. Abrupt)
- Hospitals lack a data-driven strategy



Primary Research Question

Is a Single Unsupervised Detector Sufficient?

For clinical data drift monitoring, can we rely on one approach, or do we need a hybrid strategy?

1 Implement Leak-Free Pipeline

2 Compare OCSVM vs. IF

3 Assess Hybrid Need

Methodology

PIMA Indian Diabetes Dataset Overview

768
Total Patients

9
Clinical Features

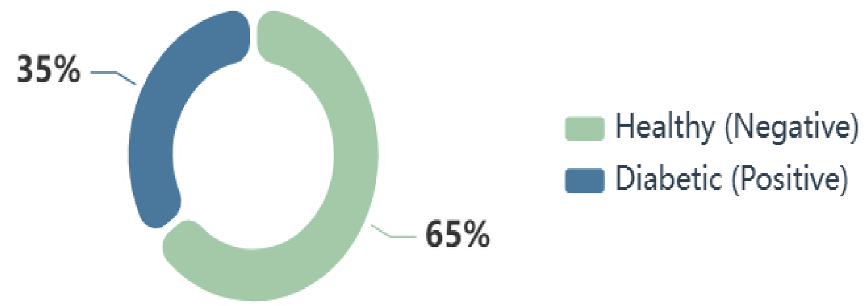
2
Classes

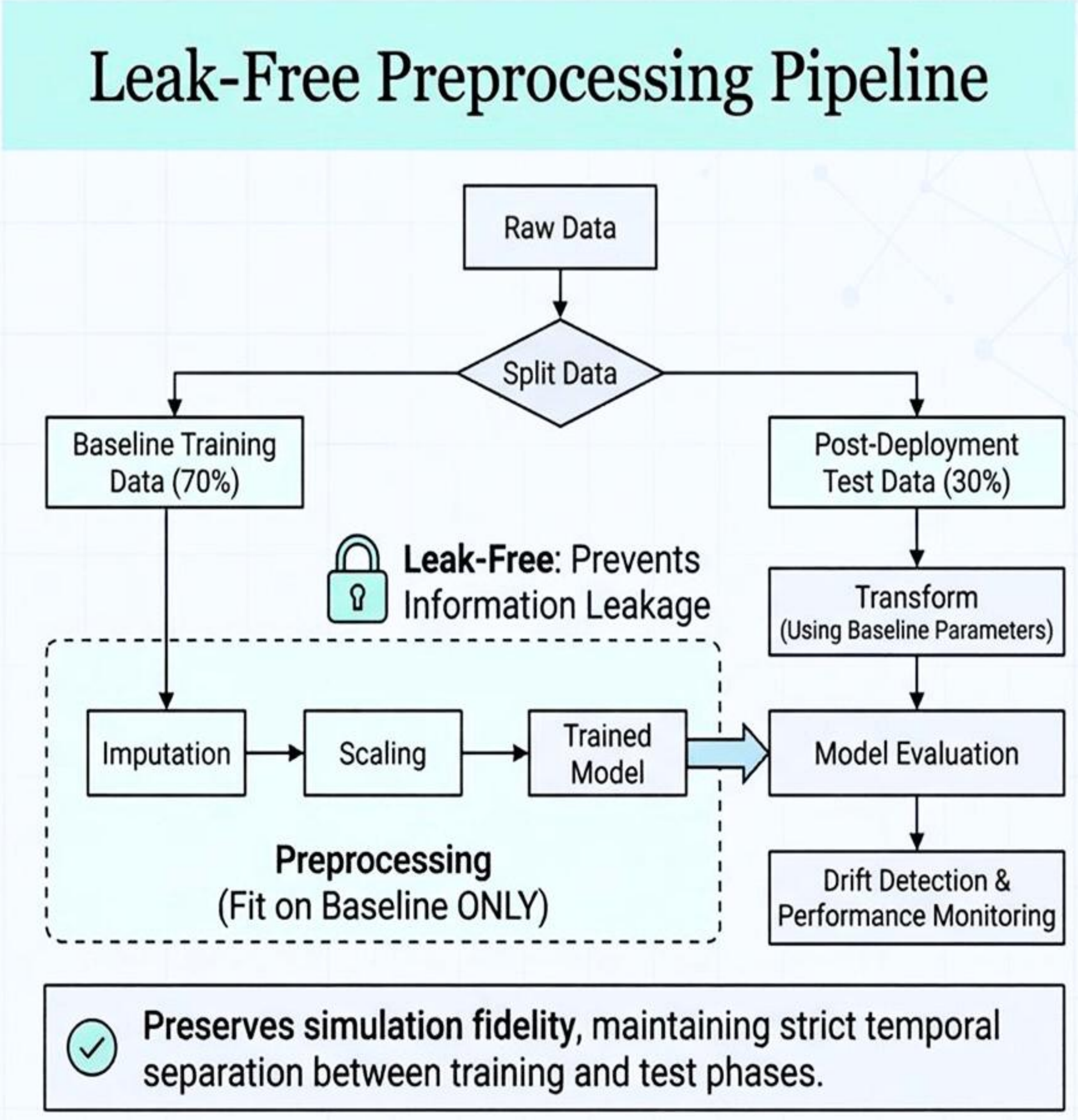
Data Split Strategy

70% Baseline (n=537)

30% Test (n=231)

Class Distribution



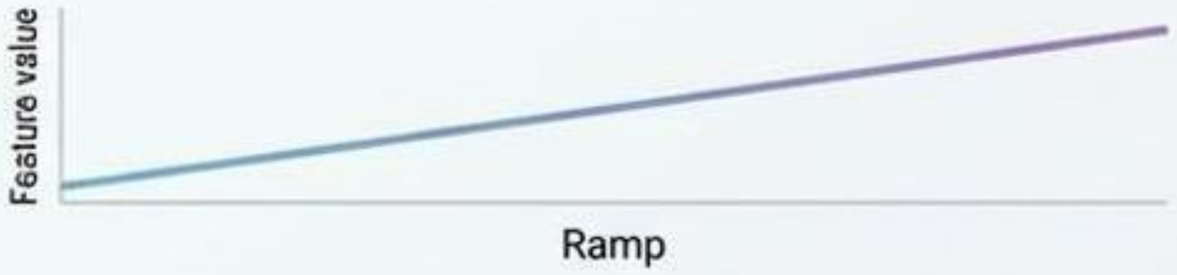


Simulated Drift Patterns: Gradual vs. Abrupt

- After establishing pipeline –

Gradual Multiplicative Drift

- **Description:** Linear ramp across 231 test samples.
- **Magnitudes:** 10%, 25%, and 40%.
- **Application:** Applied univariately & multivariately.
- **Real-world Analogy:** Equipment updates, calibration changes.



Abrupt Affine Drift

- **Description:** Instant remapping, step change.
- **Parameters:** Shift factor 0.4, Range factor 1.5.
- **Constraint:** Keeps physiologically valid ranges.
- **Real-world Analogy:** Sudden demographic shifts, policy changes.

Drift was applied to 03 key features (Glucose , BMI and Age)

Unbiased Drift Detection Lab

Experimental Setup: Model Configuration & Standardization

1. Optimized Architecture

Isolation Forest

Fixed at 100 Estimators and 256 Samples (selected via Grid Search for stability)

One-Class SVM

Utilized RBF Kernel with Gamma=0.1 to balance sensitivity against false positives

2. Fairness Controls (Explicitly Matched Thresholds)

Method

OCSVM (Nu) and IF (Contamination) parameters were numerically equated to remove bias

Nu = Contamination

Gradual Drift

Both models locked to 5% (0.05) baseline error rate

Abrupt Drift

Both models locked to 20% (0.20) baseline error rate

How We Measured Detection Performance

1 Detection Ratio

PRIMARY METRIC

DR = Outlier (Drifted) ÷ Outlier (Baseline)

Example: IF found 9.6% vs 2.2% baseline

DR = 4.40x

Results:

Stronger detection signal

2 Statistical Validation

K-S TEST

2-sample Kolmogorov-Smirnov test

Null hypothesis: Same distribution

Threshold: $p < 0.05$

Results: $p < 0.0001$

Drift confirmed beyond doubt

3 Interpretive Validation

SHAP

Shapley Additive Explanations

Purpose: Verify model logic

Shows feature importance shift

Results:

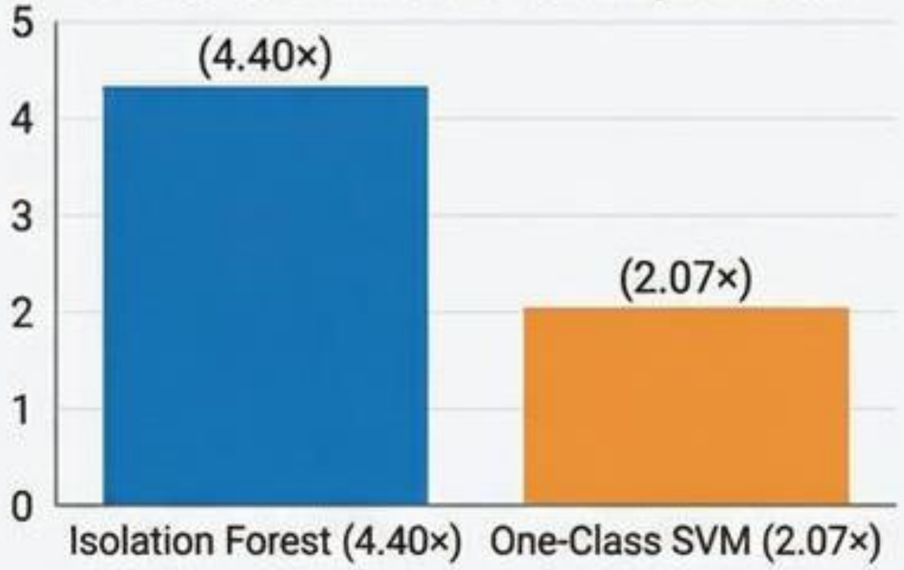
Confirming model detects correct features

- SHAP explains the anomaly score, not a ground-truth diagnosis
- Reflects unsupervised detection rather than supervised classification

Results : Gradual Drift

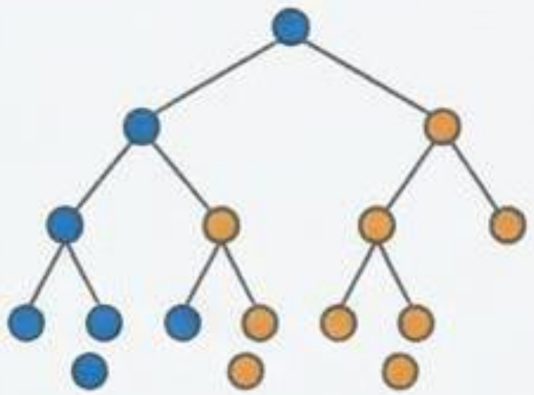
Quantitative Comparison (Under 40% Gradual Drift)					
	Algorithm	Initial Outlier Rate	Drifted Outlier Rate	Detection Ratio (Drifted/Initial)	Winner
1	Isolation Forest	2.2%	9.6%	4.40×	✓
2	One-Class SVM	11.7%	24.2%	2.07×	✗

Detection Ratio Comparison



Mechanistic Explanation: Sensitivity Advantage

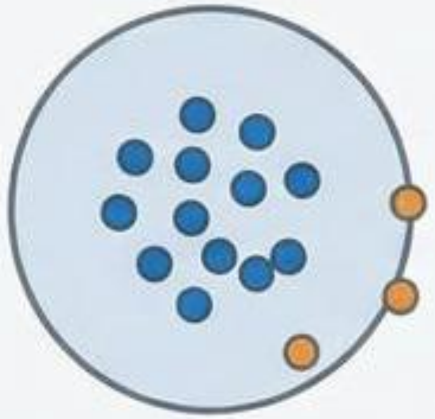
Isolation Forest (Superior Sensitivity)



Sparse drifting points are rapidly isolated by few tree splits, leading to high detection efficiency.

✓ 2.1-fold sensitivity advantage over SVM.

One-Class SVM (Weak Signal)



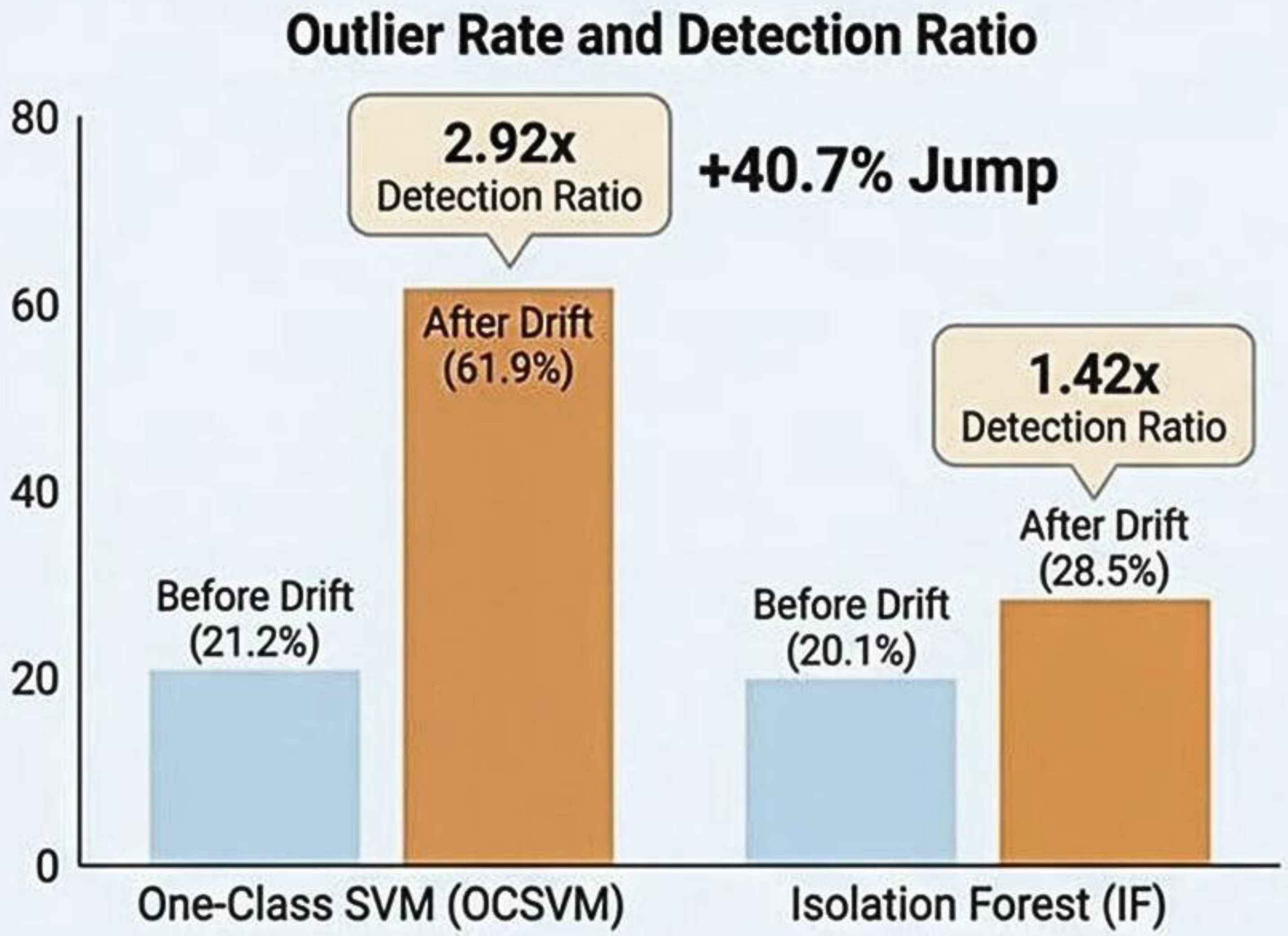
The same slow shift remains largely inside the SVM boundary, producing a weak signal for detection.

✗ Limited sensitivity to gradual shifts.

Key Insight: Isolation Forest demonstrates significantly higher sensitivity (4.40× vs 2.07×) to gradual drift due to its ability to rapidly isolate sparse drifting points, while SVM struggles to differentiate slow shifts from normal data.

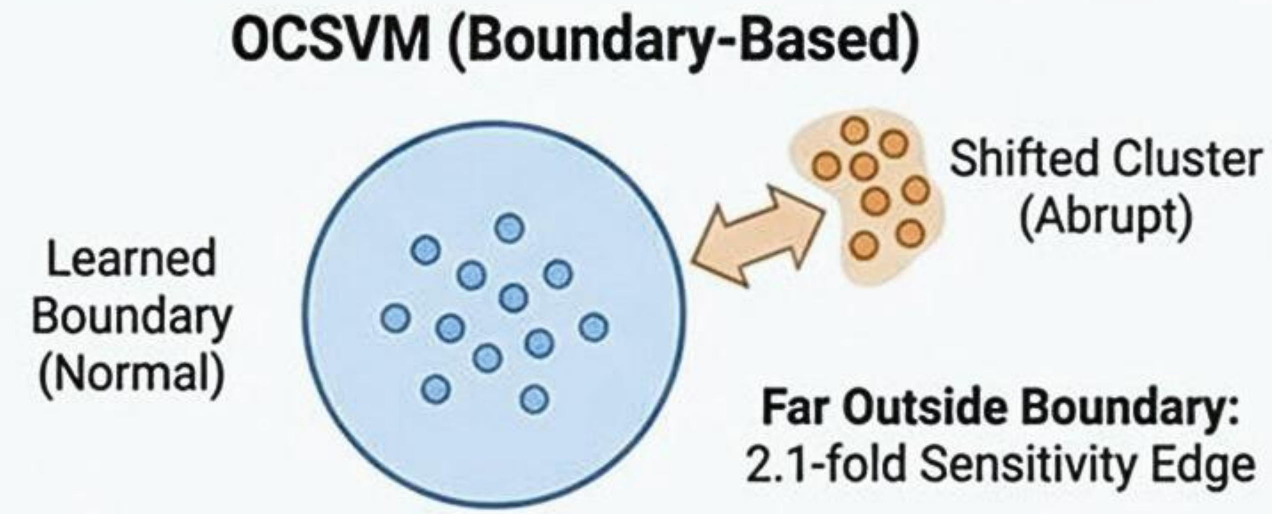
Results : Abrupt Drift

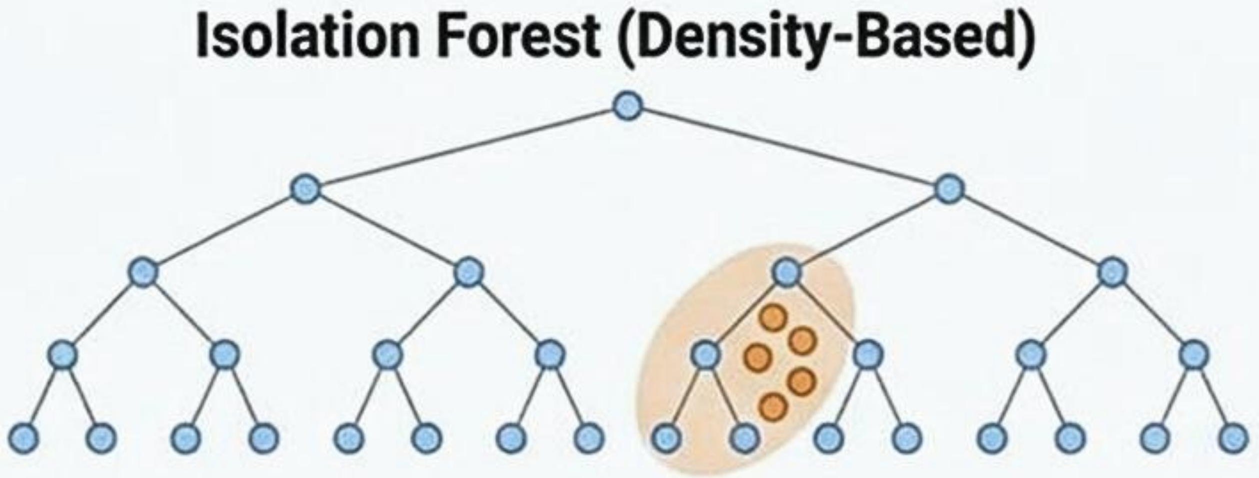
Performance Comparison: Abrupt Affine Drift



Model	Before %	After %	Ratio	Status
OCSVM	21.2%	61.9%	2.92x	✅
IF	20.1%	28.5%	1.42x	⚠️

Mechanistic Insight: Why OCSVM Outperforms



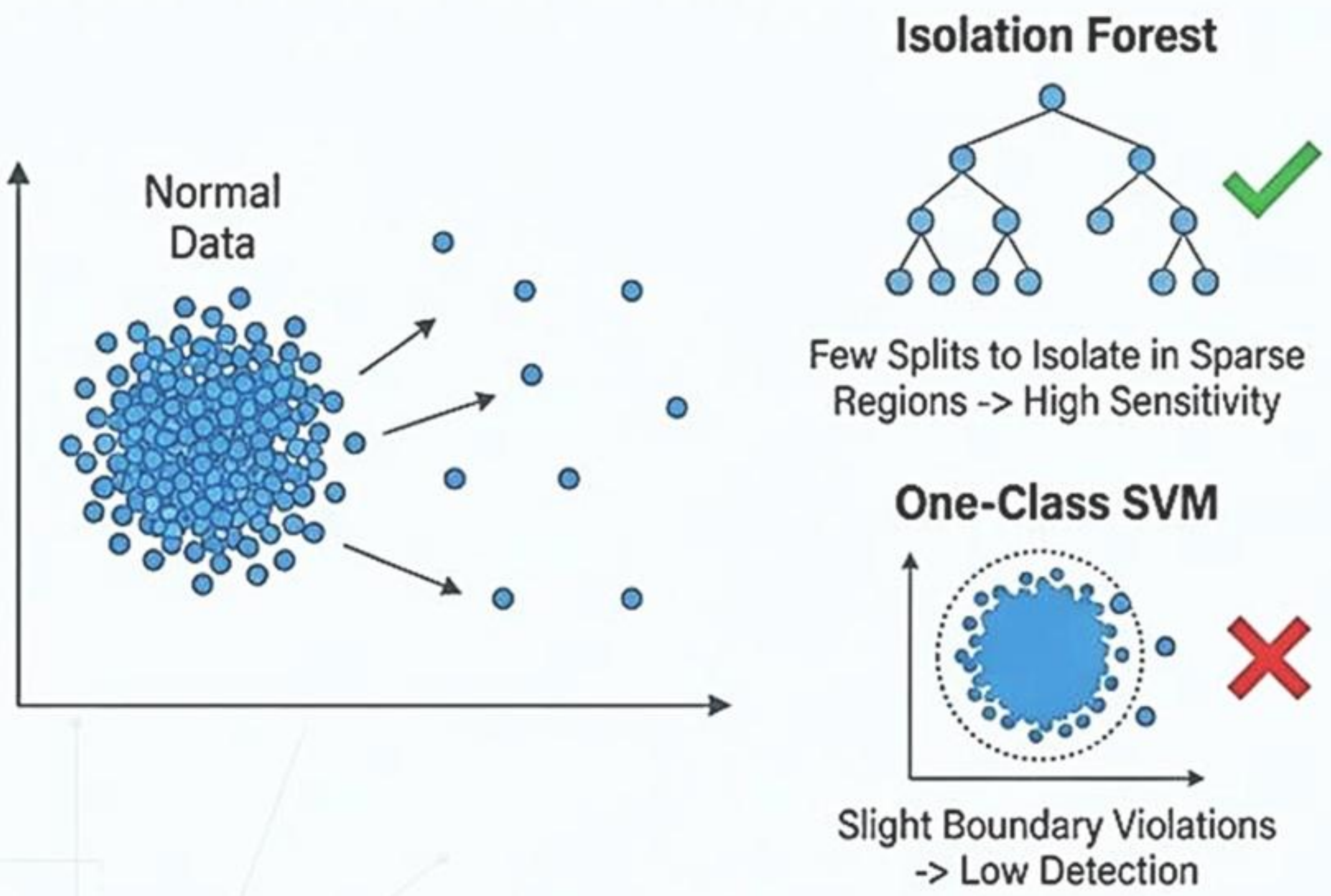


Maintained Density: Looks Normal to Partition Trees, Lower Sensitivity

Complementary Algorithmic Strengths: OCSVM excels at detecting global shifts, while IF is robust to local anomalies.

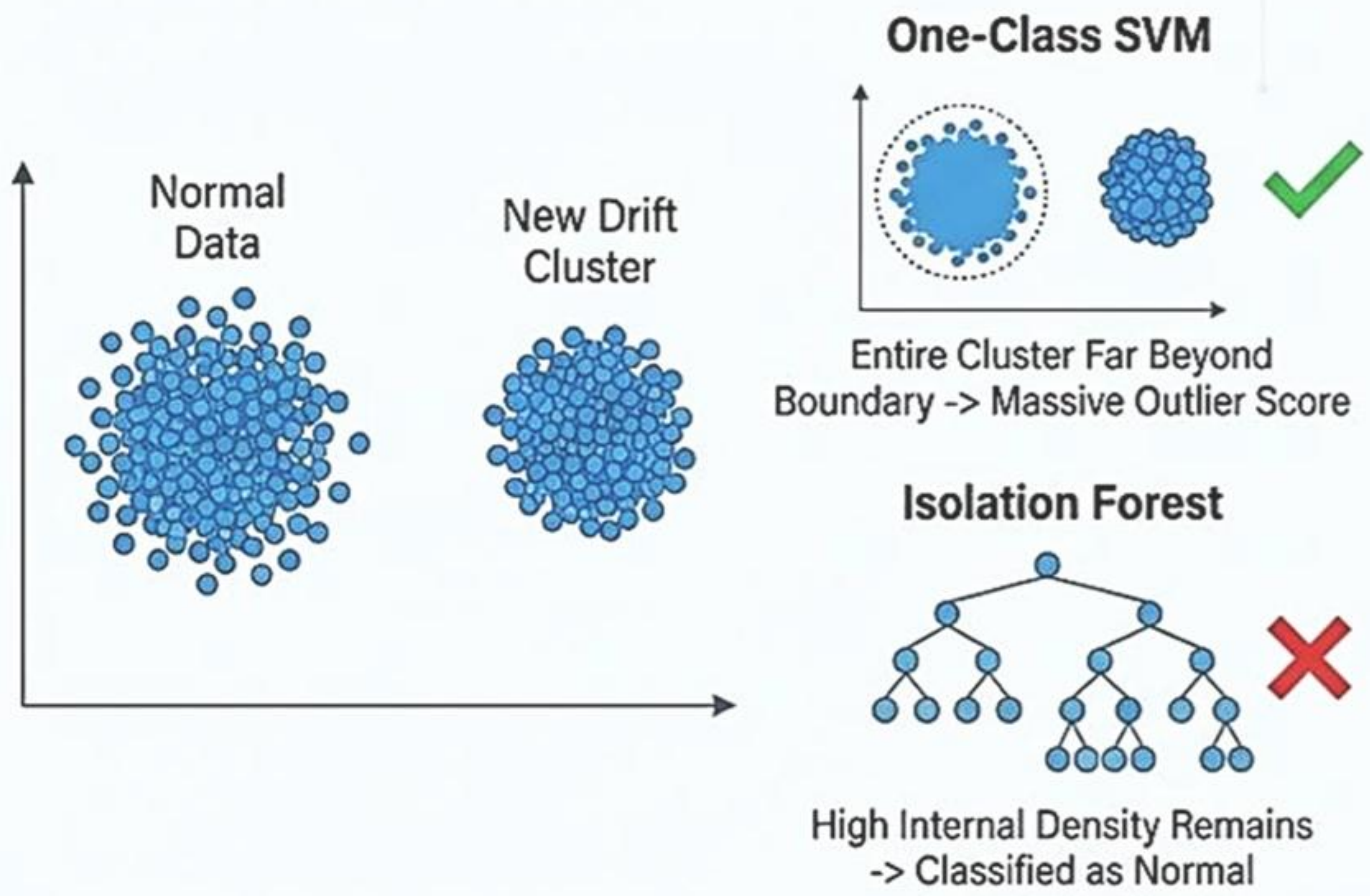
Mechanistic Explanation: The Geometric Divergence

GRADUAL DRIFT (Sensitivity to Sparsity)



Gradual drift pushes points into sparse space, amplifying Isolation Forest's sensitivity as few splits are needed, while One-Class SVM registers minor deviations.

ABRUPT DRIFT (Sensitivity to Location)



Abrupt drift relocates an entire cluster; One-Class SVM detects large boundary violation, while Isolation Forest struggles with the new cluster's high internal density.

Conclusion : The Trade-off Truth

Strengths of these two models are perfectly inverse.

Algorithm	Gradual Drift Scenarios	Abrupt Shifts
Isolation Forest (IF)	✓ **Excels.** Demonstrates superior sensitivity to slow, accumulating shifts.	✗ **Struggles.** Less effective at detecting sudden, dramatic changes.
One-Class SVM (OCSVM)	✗ **Struggles.** Often misses gradual, non-linear patterns.	✓ **Dominates.** Highly effective at identifying sudden, stark anomalies.

This proves that relying on a single unsupervised model creates a dangerous 'blind spot'.

We need a hybrid system to ensure patient safety.

✓ Excels / Dominates for specific scenarios

✗ Struggles / Misses for alternate scenarios

Study Limitations

Four key constraints impact the broader application of current findings.

1. Synthetic Drift vs. Real Patterns

- Synthetic drift models lack complex temporal dependencies found in real longitudinal hospital data.

2. Limited Test Set Variability

- Small test set (N=231) may under-represent variability present in large institutional databases.

3. Low-Dimensional Feature Space

- 8-feature model differs from high-resolution medical imaging where drift is nonlinear.

4. Unmonitored Concept Drift

- Algorithms monitor input features only, overlooking shifts in decision boundaries between features and outcomes.

Future Work & Recommendations

① Real-World Validation

- Validate findings on authentic hospital datasets.
- Focus on data containing genuine temporal drift.

② Method Comparison

- Compare unsupervised methods.
- Benchmark against supervised detectors (DDM, ADWIN) when labels exist.

③ Hybrid Ensemble

- Engineer a hybrid ensemble.
- Run One-Class SVM and Isolation Forest in parallel for full-spectrum coverage.

Immediate Practical Recommendation

✅ Deploy both detectors concurrently in hospitals. Ensure neither gradual nor abrupt drift escapes surveillance.

THANK YOU