Drift Detection Metrics – Explanation & Justification

AI Model Governance & Monitoring Lab

---

1. What is Data Drift?

Data drift is the change in the distribution of input data over time. In production machine learning, this lead to "silent failure" where the model remains technically operational but its predictive accuracy degrades because the real-world data no longer matches the training data.

2. Metrics Implemented

Population Stability Index (PSI)

PSI measures how much a variable has shifted in distribution between two points in time (Reference vs. Production).

- **Why suitable:** It is scale-invariant and widely used in financial industry to detect covariate shift.

PSI Value	Interpretation
< 0.1	No significant shift; model is stable.
0.1 – 0.25	Moderate shift; investigate data sources.
> 0.25	Significant shift; actions required (Retraining).

Kolmogorov–Smirnov (KS Test)

A non-parametric statistical test that compares the cumulative distributions of two data samples.

- **Why suitable:** It detects differences in location and shape of empirical distributions without assuming normality.
- **Interpretation:** A p-value < 0.05 indicates we reject the null hypothesis that distributions are the same (Drift Detected).

3. Key Observations

In our simulation, increasing the average 'Income' by 50% resulted in a **PSI of ~0.65**, which correctly triggered a "Significant Shift" alert. This correlated with a 12% drop in F1-score performance.