Model Drift

1️⃣ What is Model Drift?

In Machine Learning, we **assume the future resembles the past**.

When this assumption breaks, **model performance degrades over time** because the **real world has changed but the model has not**.

This phenomenon is called **Model Drift**.

Formally:

- The model was trained on historical data
- The production data distribution changes
- Predictions become unreliable

2️⃣ Types of Model Drift

🔷 1. Data Drift (Covariate Shift)

What changes: Input features

What stays same: Relationship between X and Y

Mathematical View

$$P(X_{train}) \neq P(X_{prod}), \quad P(Y|X) \text{ remains constant}$$

Intuition

- Model sees **new kinds of data**
- It has never learned how to behave there

Example

- Face recognition trained on **young faces**

- Used years later on **aged faces**
- Labels are correct, inputs changed

🔷 2. Concept Drift

What changes: Relationship between X and Y

Data may look similar, but meaning changes

Mathematical View

$$P(Y|X) \text{ changes}$$

Intuition

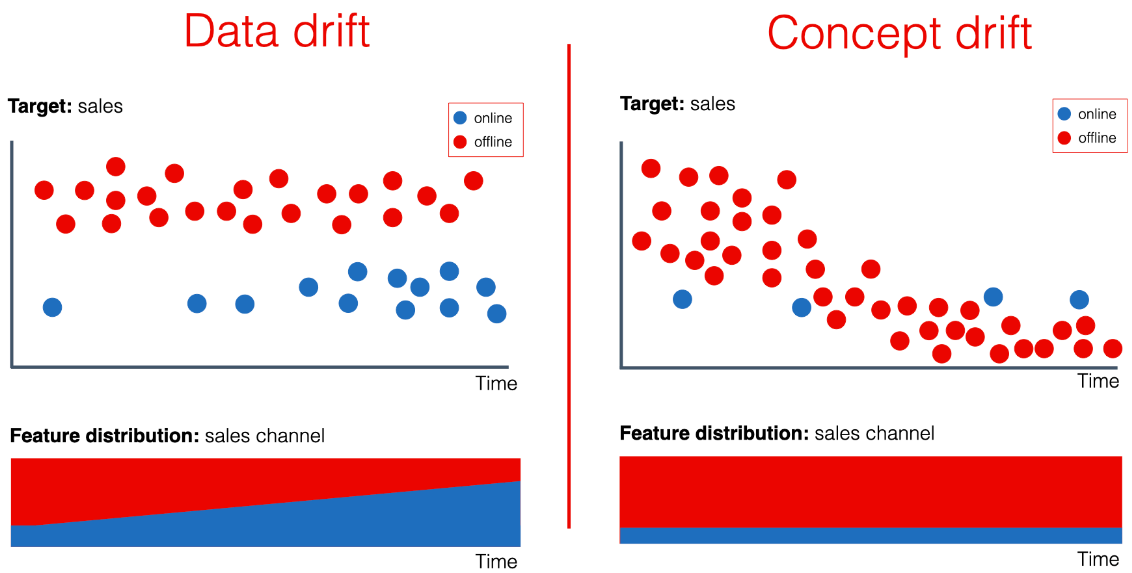
- Old decision rules are no longer valid

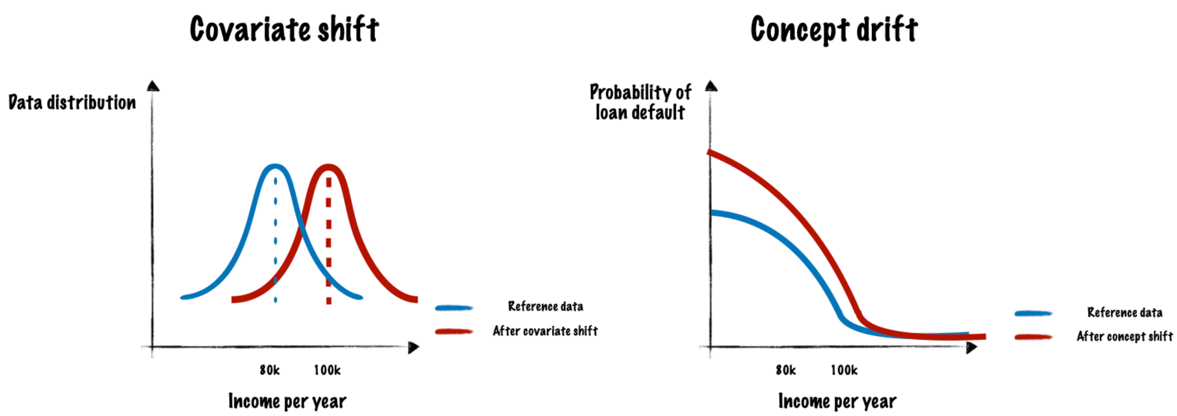
Example

- Spam detection:
  - 2010 → “Win an iPad”
  - 2024 → “Urgent: Invoice Attached”

Same words, **different intent**.

3️⃣ Data Drift vs Concept Drift (Visual Intuition)





Aspect	Data Drift	Concept Drift
What moves?	Data points (X)	Decision boundary
Model issue	Never saw this data	Rule is outdated
Analogy	Using Paris map in New York	Using 1700s map in Paris

4️⃣ Drift Behavior Over Time

Types Based on Change Pattern

Type	Description	Example
Sudden Drift	Abrupt change	COVID demand spike
Gradual Drift	Slow evolution	Inflation over years
Recurring Drift	Cyclical	Seasonal sales

5️⃣ Drift Detection — Mathematical Tools

🔷 1. Kolmogorov–Smirnov (KS) Test

Purpose: Compare two distributions

Used for: Numerical features

Hypothesis

- H_0: Both samples come from same distribution
- Reject if **p-value < 0.05**

🔷 2. Population Stability Index (PSI)

Used heavily in **finance & credit scoring**.

$$PSI = \sum (Actual\% - Expected\%) \times \ln\left(\frac{Actual\%}{Expected\%}\right)$$

Interpretation

- PSI < 0.1 → No drift
- 0.1–0.2 → Moderate drift
- ≥ 0.2 → Significant drift (action required)

🔷 3. Page-Hinkley Test

- Online / streaming drift detection
- Detects **sudden mean shifts**

- Used in real-time systems

6️⃣ Hands-On Python: Data Drift Detection (KS Test)

📌 What This Code Does

- Simulates **training vs production data**
- Visualizes distribution change
- Detects drift statistically

🧠 Python Code

```
import numpyas np
import matplotlib.pyplotas plt
import seabornas sns
from scipy.statsimport ks_2samp

# 1. Training Data (Baseline)
np.random.seed(42)
train_data = np.random.normal(loc=0, scale=1, size=1000)

# 2. Production Data (Drifted)
prod_data = np.random.normal(loc=0.5, scale=1, size=1000)

# 3. Visualize Distributions
plt.figure(figsize=(10,5))
sns.kdeplot(train_data, label="Training Data", fill=True)
sns.kdeplot(prod_data, label="Production Data", fill=True)
plt.title("Data Drift Visualization")
plt.legend()
plt.show()

# 4. KS Test
statistic, p_value = ks_2samp(train_data, prod_data)
```

```
print(f"KS Statistic: {statistic:.4f}")
print(f"P-Value: {p_value:.4e}")

# 5. Interpretation
if p_value <0.05:
print("🚨 DRIFT DETECTED: Distributions differ significantly.")
else:
print("✅ NO DRIFT: Distributions are similar.")
```

📊 How to Read Output

- **Very small p-value (e.g. 1e-10)**

  → Almost zero chance distributions are same

  → **Confirmed data drift**
