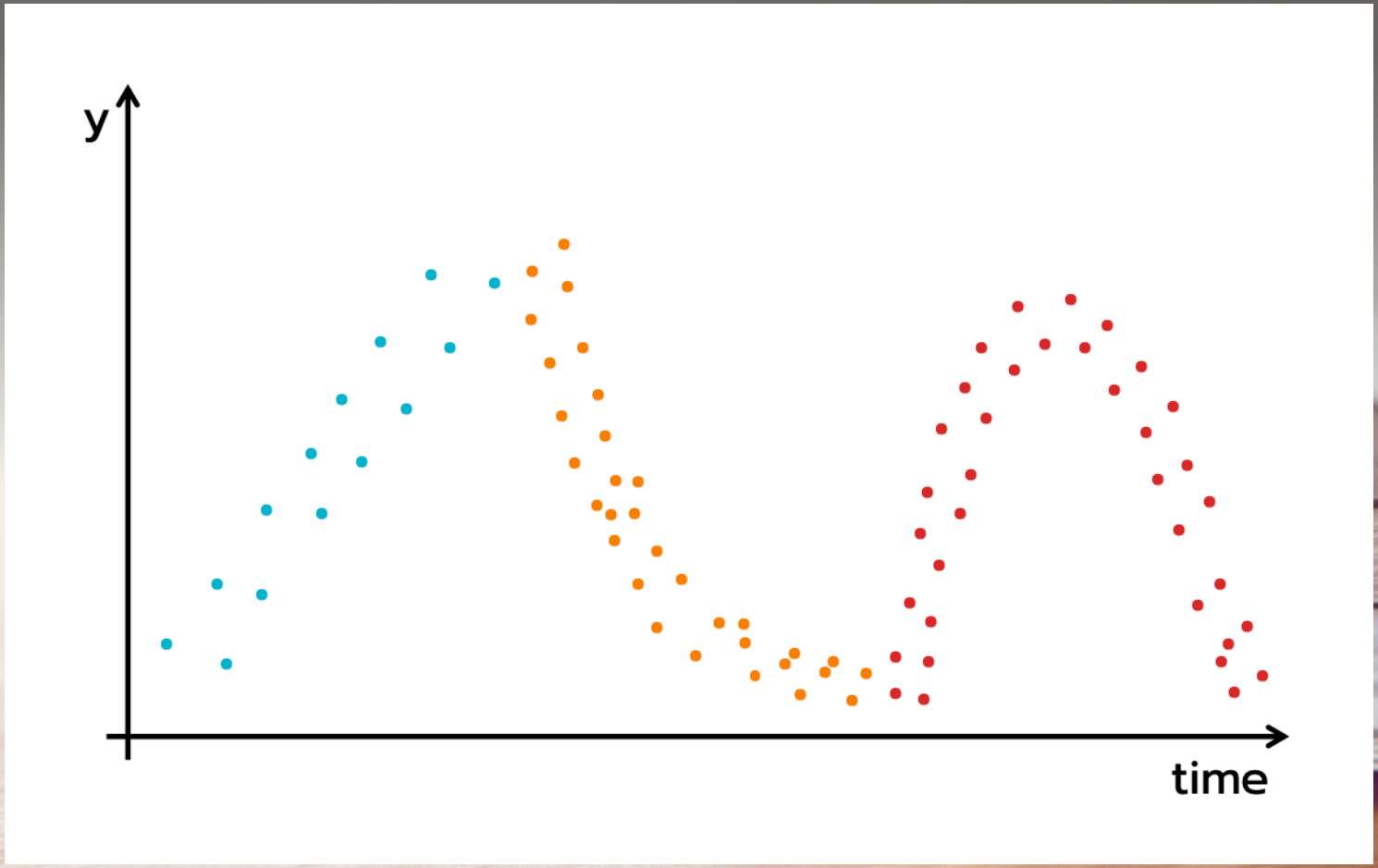

CONCEPT
DRIFT

CONCEPT DRIFT

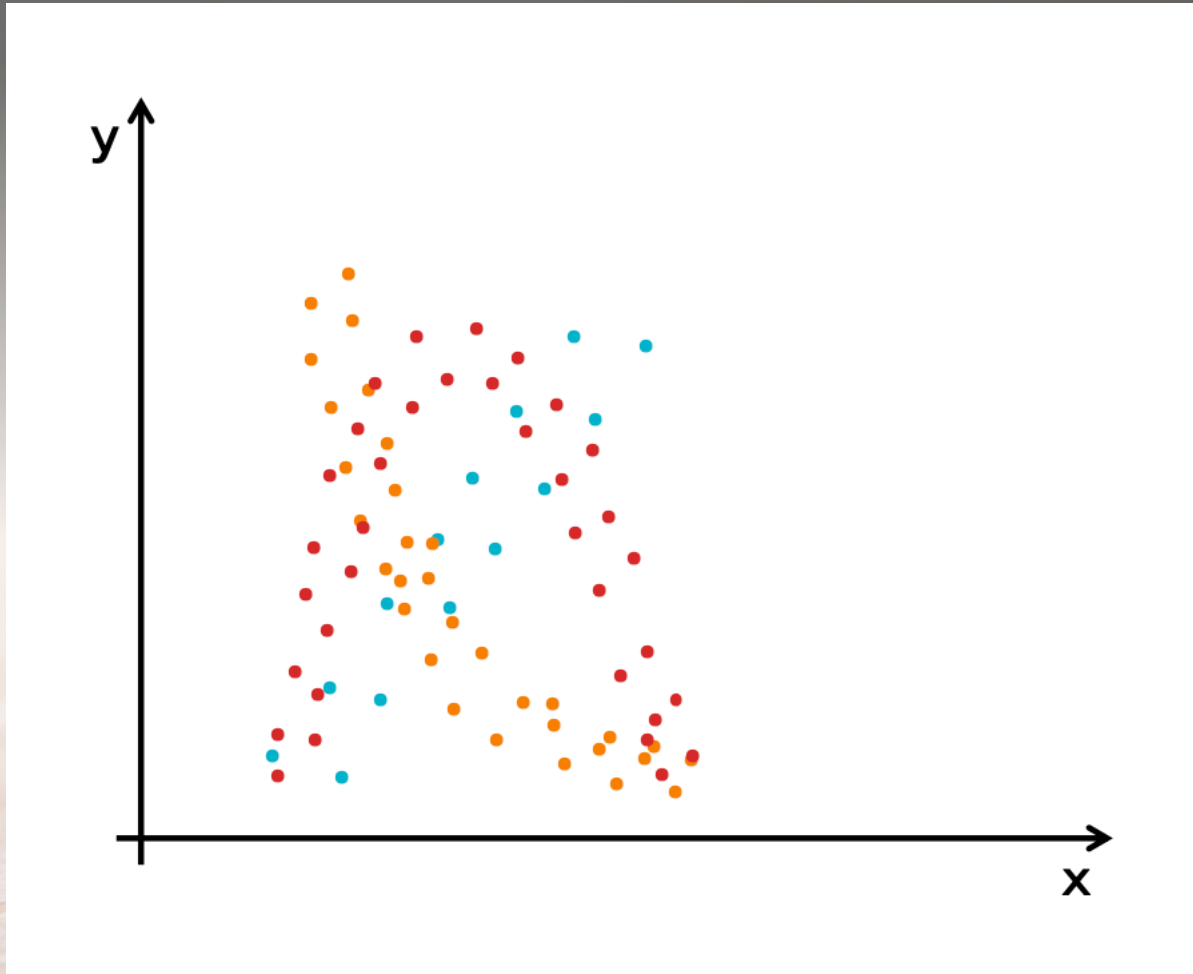
- Introduction
- Overview of concept drift
- Before we start
- Concept drift with SL
- Concept drift with UL
- Concept drift with RL

Introduction

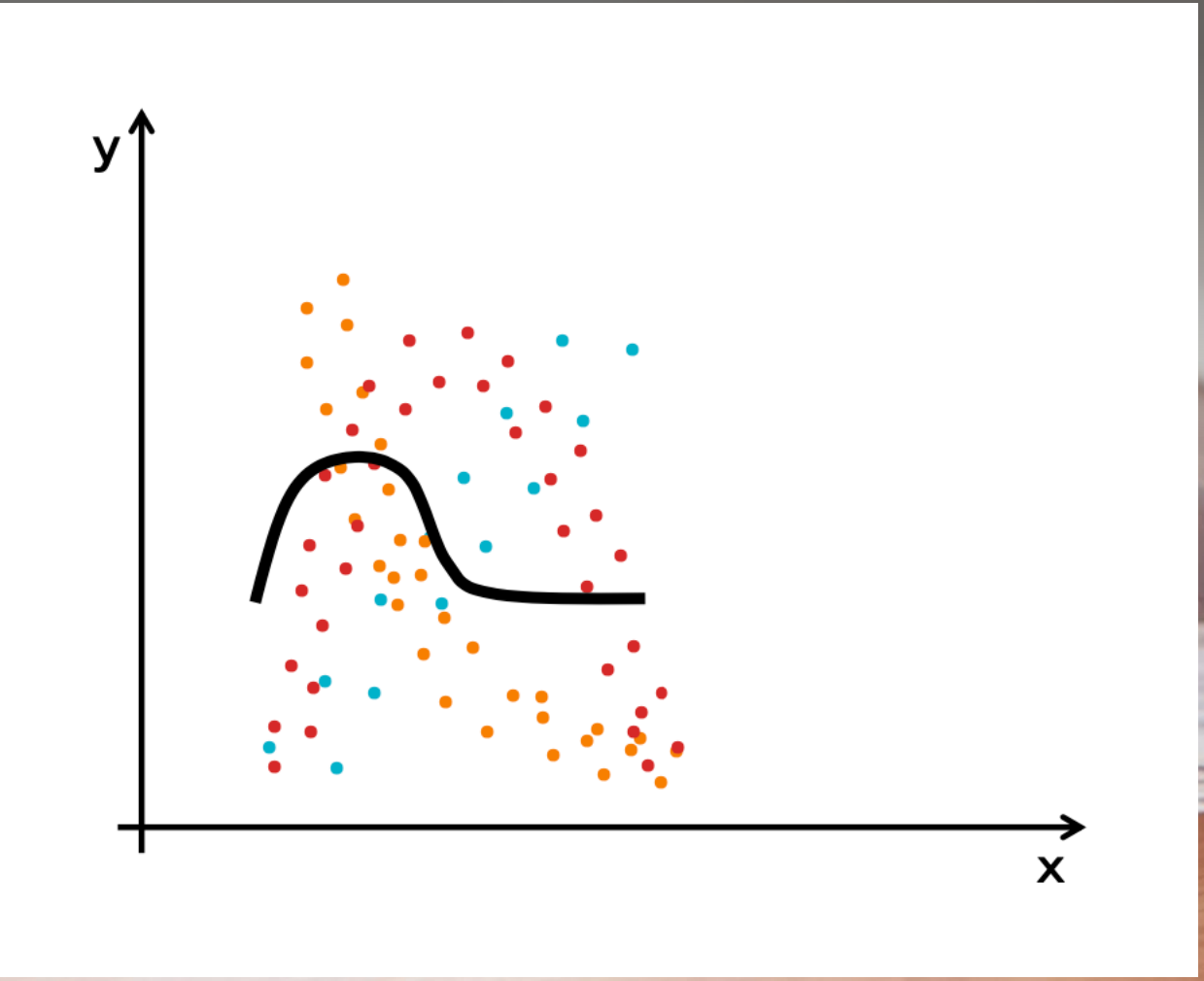
Introduction



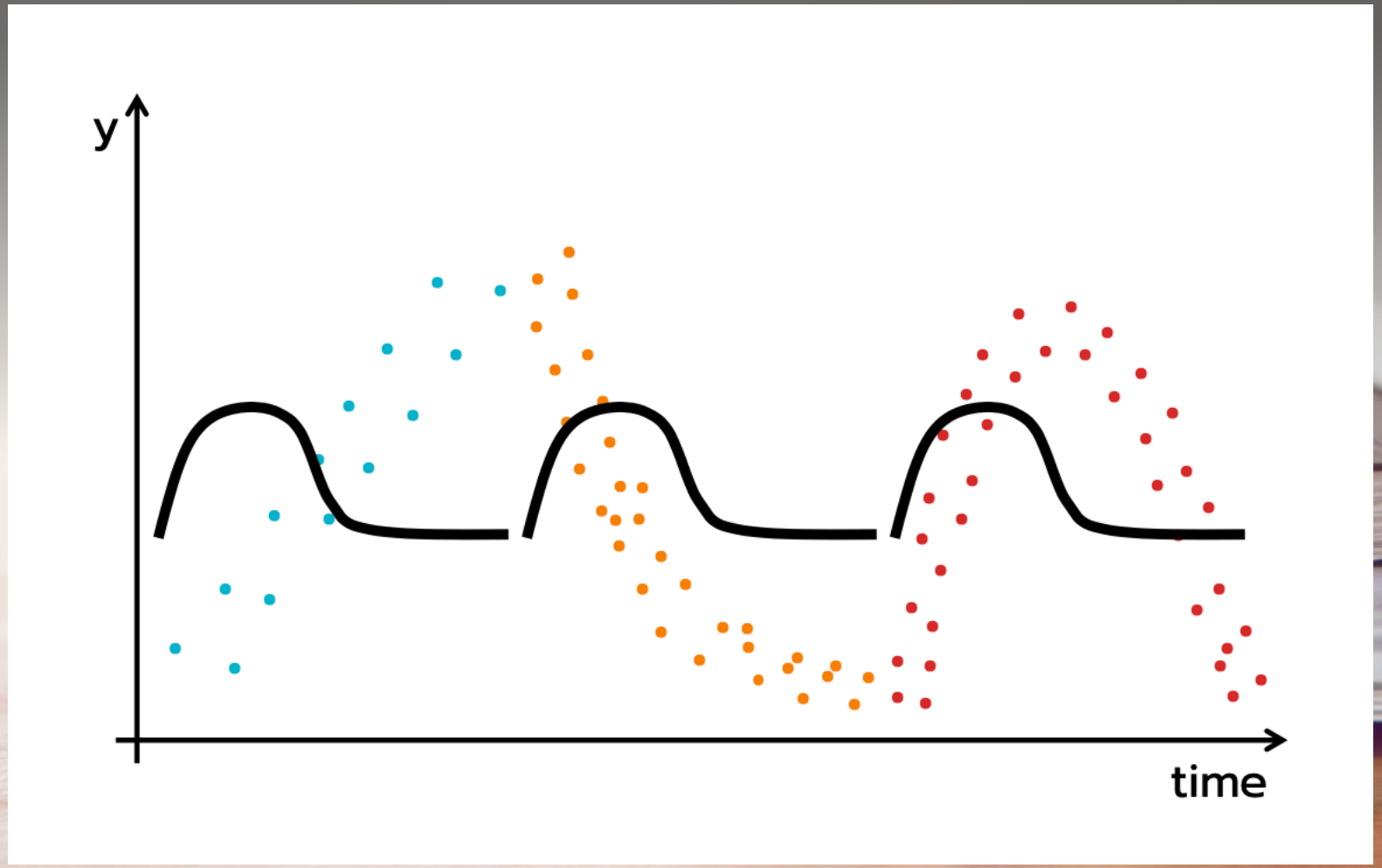
Introduction



Introduction



Introduction



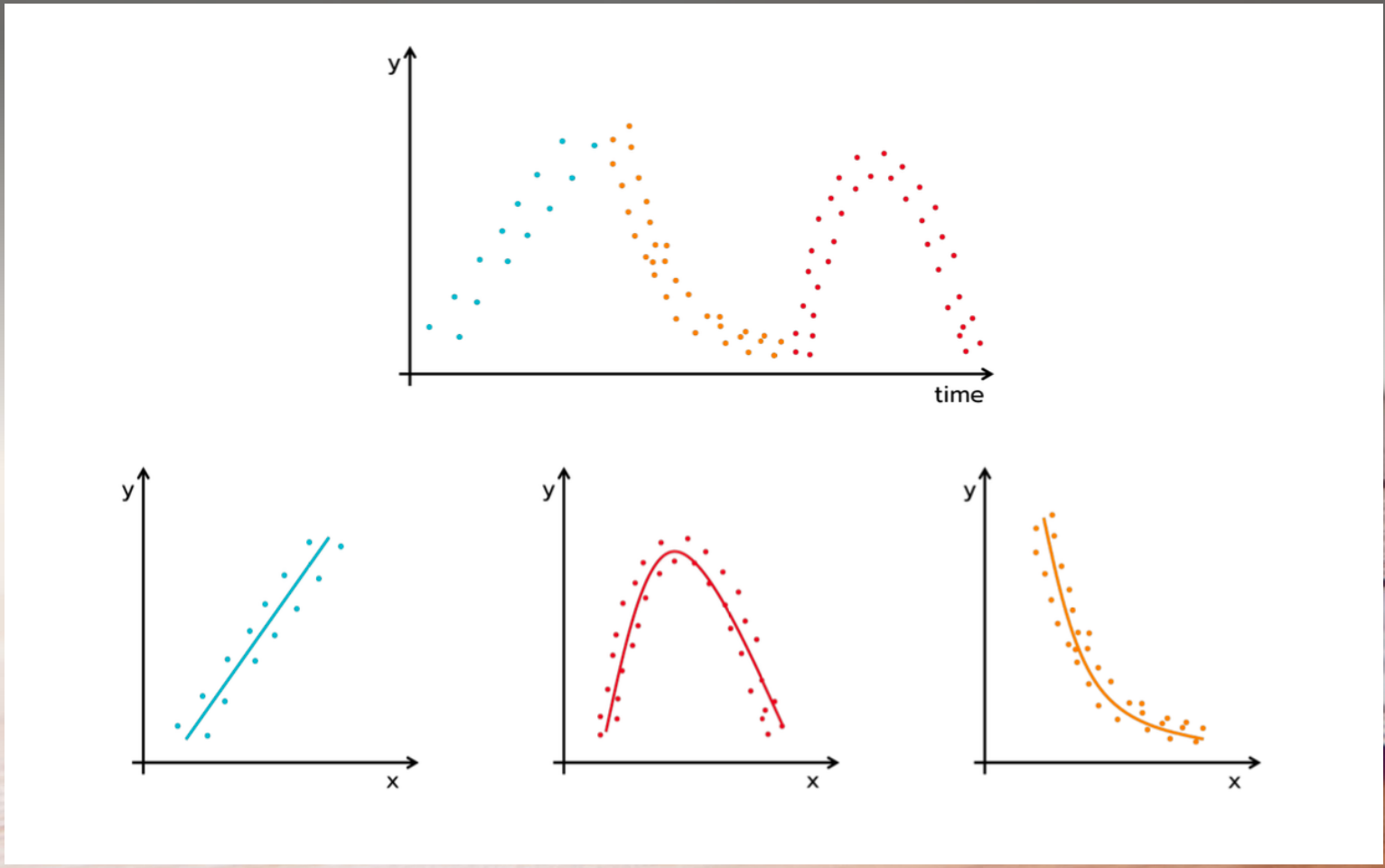
Introduction

สิ่งที่พบเห็นตามมาบ่อย ๆ คือ

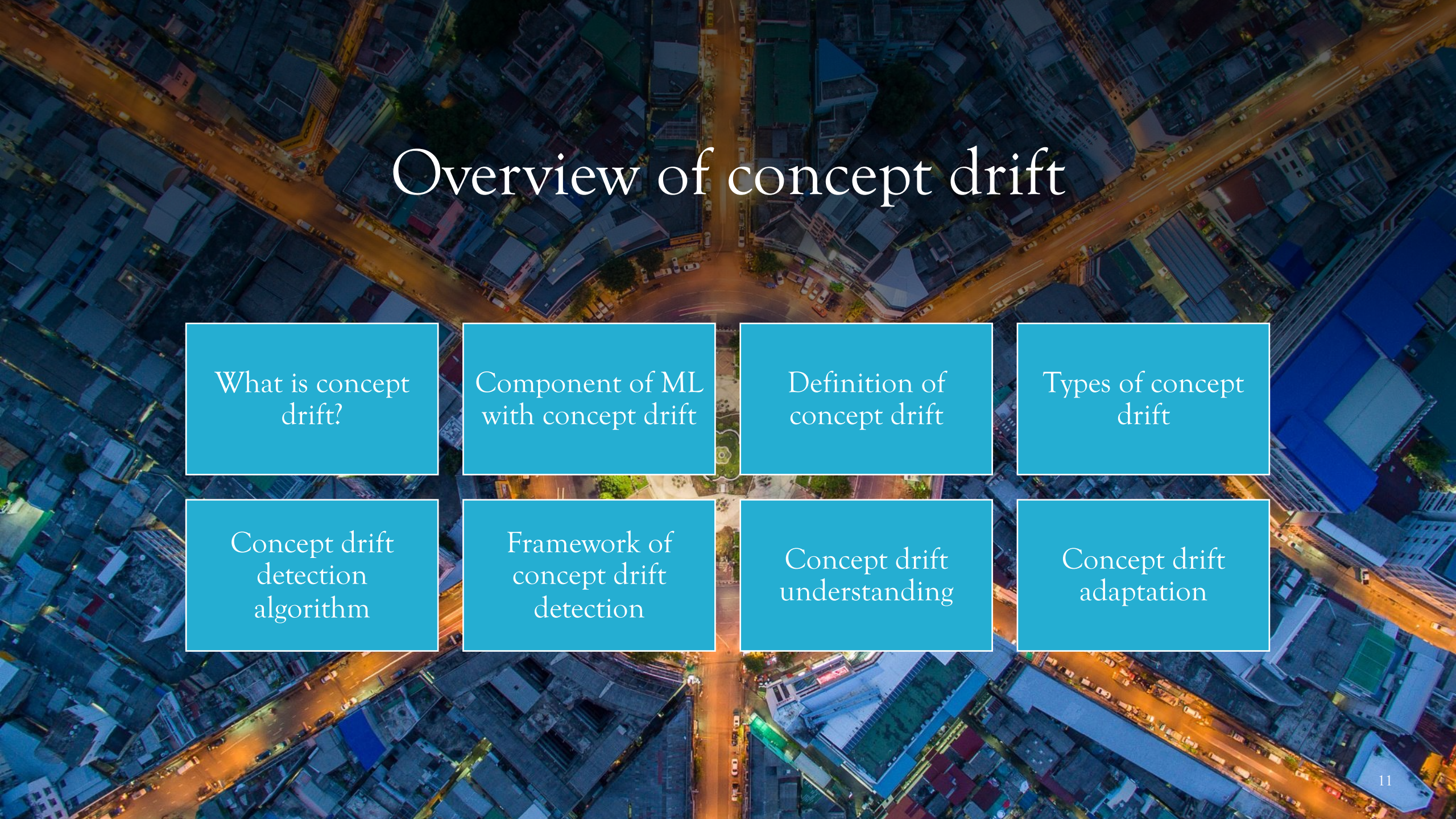
การเพิ่ม non informative feature เข้าไป เพื่อให้ ML สามารถ fit โมเดลได้ → นำไปสู่การ overfitting ในที่สุด

นี่คือหลุมพรางสำหรับหลาย ๆ คนที่เข้ามาแก้ปัญหานี้

Introduction



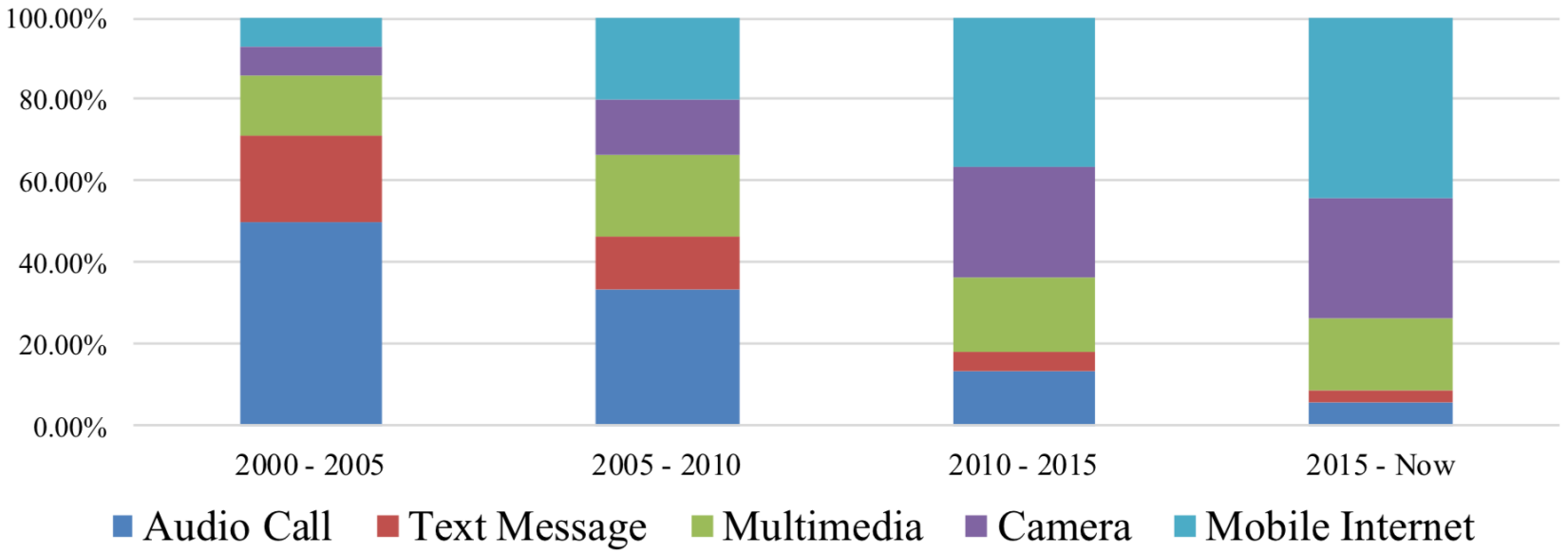
Overview of concept drift


Overview of concept drift
What is concept drift?
Component of ML with concept drift
Definition of concept drift
Types of concept drift
Concept drift detection algorithm
Framework of concept drift detection
Concept drift understanding
Concept drift adaptation

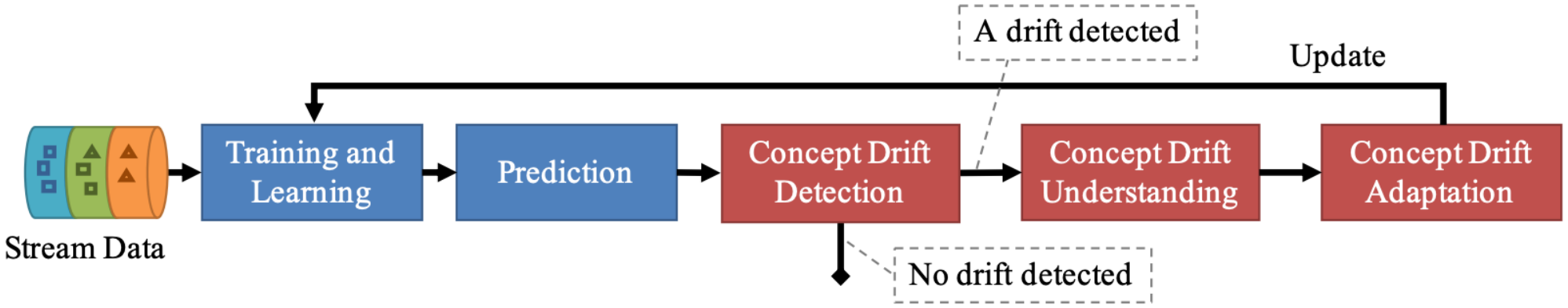
What is concept drift?

Concept drift in machine learning refers to the phenomenon where the statistical properties of the target variable, or the relationship between input and output variables, change over time. This can lead to a decline in model performance since the model was trained on historical data that no longer represents the current data distribution.

What is concept drift?



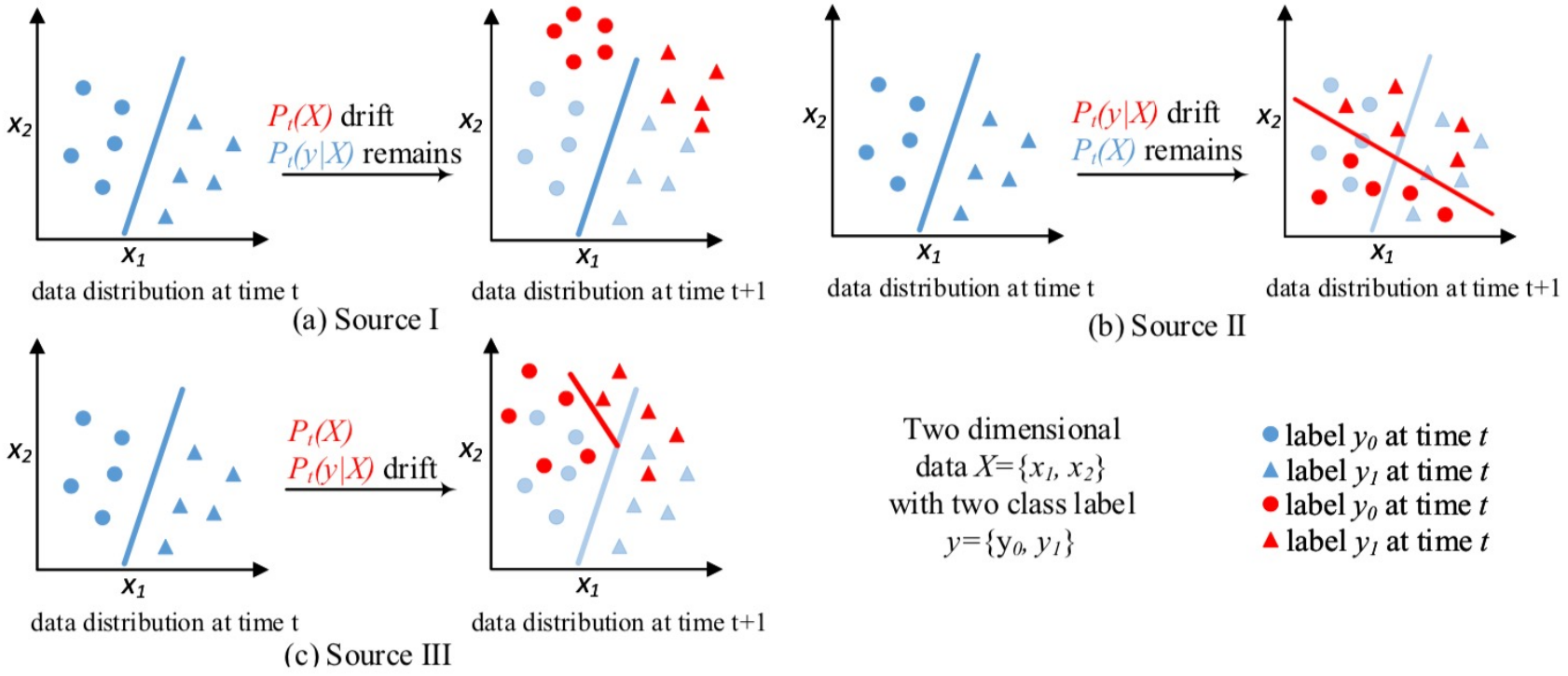
Component of Machine Learning with Concept Drift



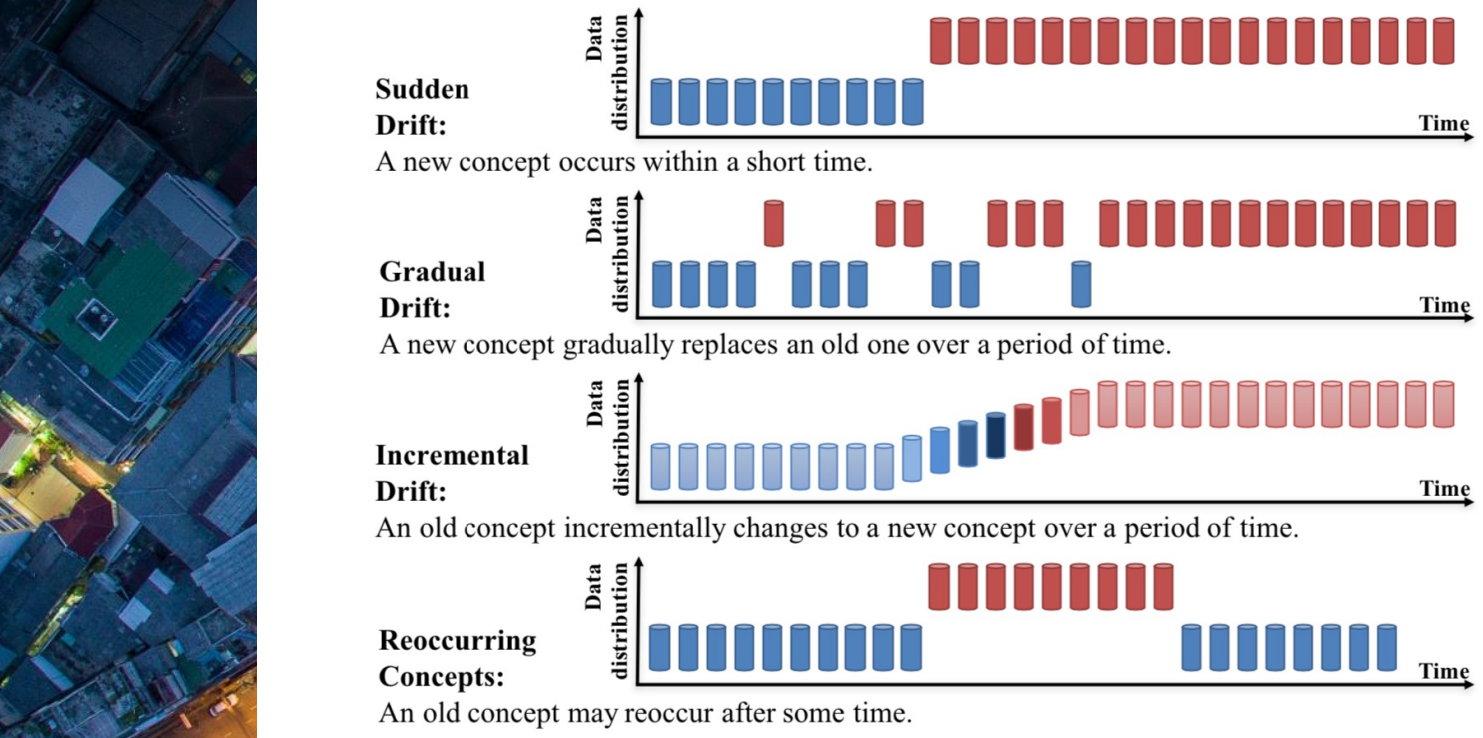
Definition of concept drift

$$P_t(X, y) = P_t(X) \times P_t(y|X)$$

Definition of concept drift



Types of concept drift



Sudden Drift:
A new concept occurs within a short time.

Gradual Drift:
A new concept gradually replaces an old one over a period of time.

Incremental Drift:
An old concept incrementally changes to a new concept over a period of time.

Reoccurring Concepts:
An old concept may reoccur after some time.

Concept drift detection

Error Rate-based Drift Detection

Data Distribution-based Drift Detection

Multiple Hypothesis Test Drift Detection

Concept drift detection

- **Error Rate-based Drift Detection**
  - Learning with drift detection
  - Exponentially weighted moving average charts for detecting concept drift
  - Detecting concept drift using statistical testing
  - Learning from time-changing data with adaptive windowing

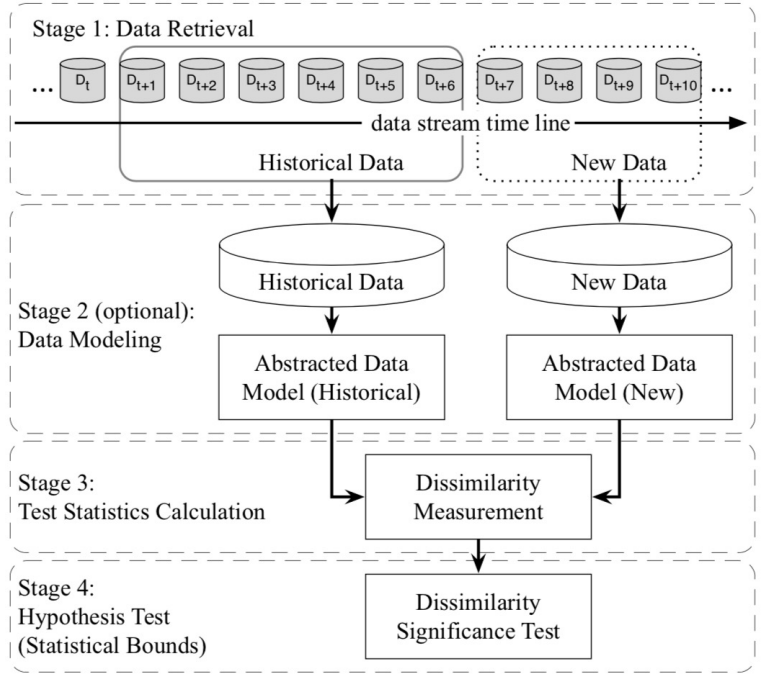
Concept drift detection

- **Data Distribution-based Drift Detection**
  - Detecting change in data streams
  - An information-theoretic approach to detecting changes in multi- dimensional data streams
  - Concept drift detection via competence models

Concept drift detection

- **Multiple Hypothesis Test Drift Detection**
  - A selective detector ensemble for concept drift detection
  - Three-layer concept drifting detection in text data streams
  - Just-in-time adaptive classifiers part i: Detecting nonstationary changes

Framework of Concept Drift Detection



Concept drift understanding

WHEN

HOW

WHERE

Concept drift understanding

How

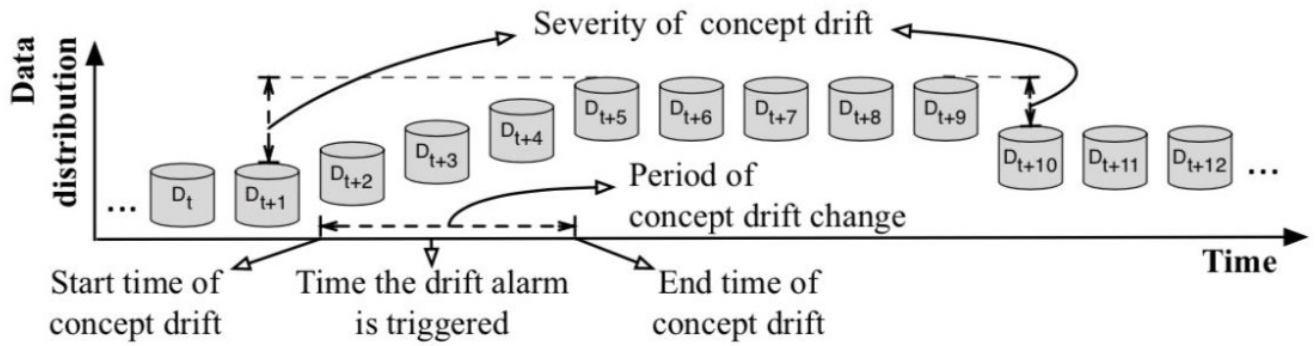
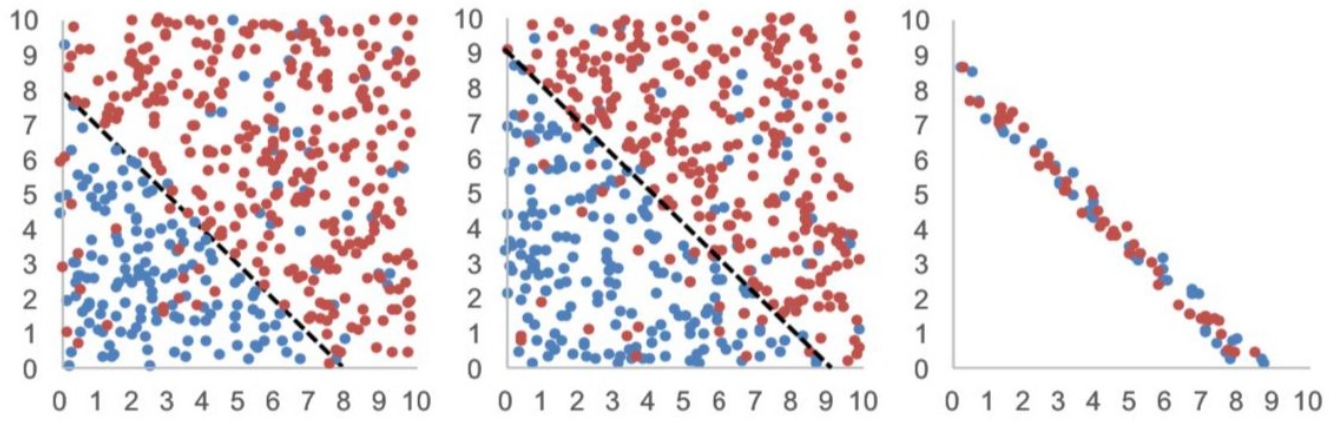


Fig. 11. An example of the occurrence time and the severity of concept drift. One incremental drift starts to change at $t+1$ and ends at $t+5$. The other sudden drift occurs between $t+9$ and $t+10$. The severity of these two concept drifts (D_{t+1}-D_{t+5} and D_{t+9}-D_{t+10}) is marked with brackets.

Concept drift understanding

Where

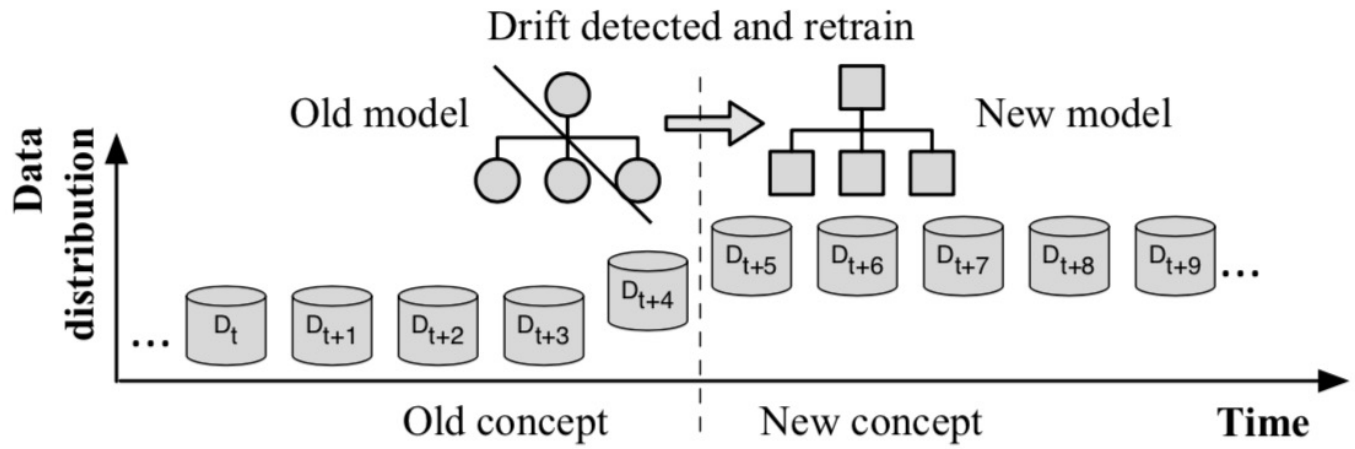


(a) Data at time t (b) Data at time $t+1$ (c) Data in drift regions

Fig. 12. An example of the drift regions of concept drift.

Concept drift adaptation

- **Retrain**



Concept drift adaptation

Adaptive Ensemble Method

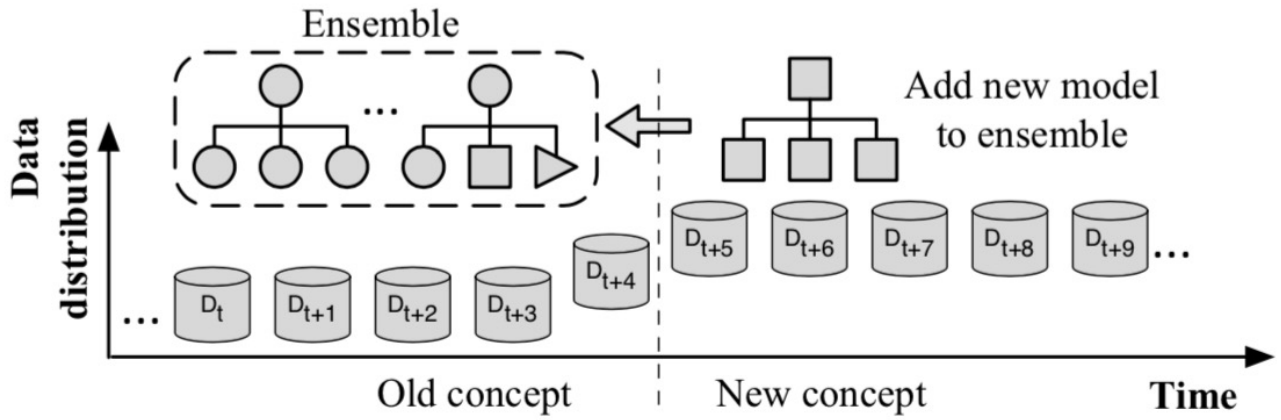


Fig. 14. A new base classifier is added to the ensemble when a concept drift occurs.

Concept drift adaptation

Partially Update Model

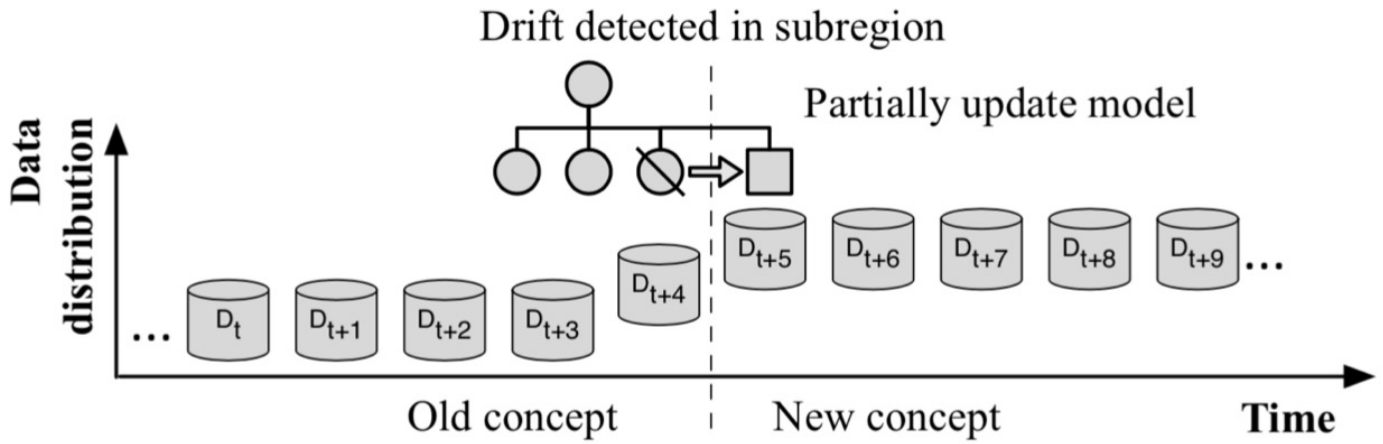
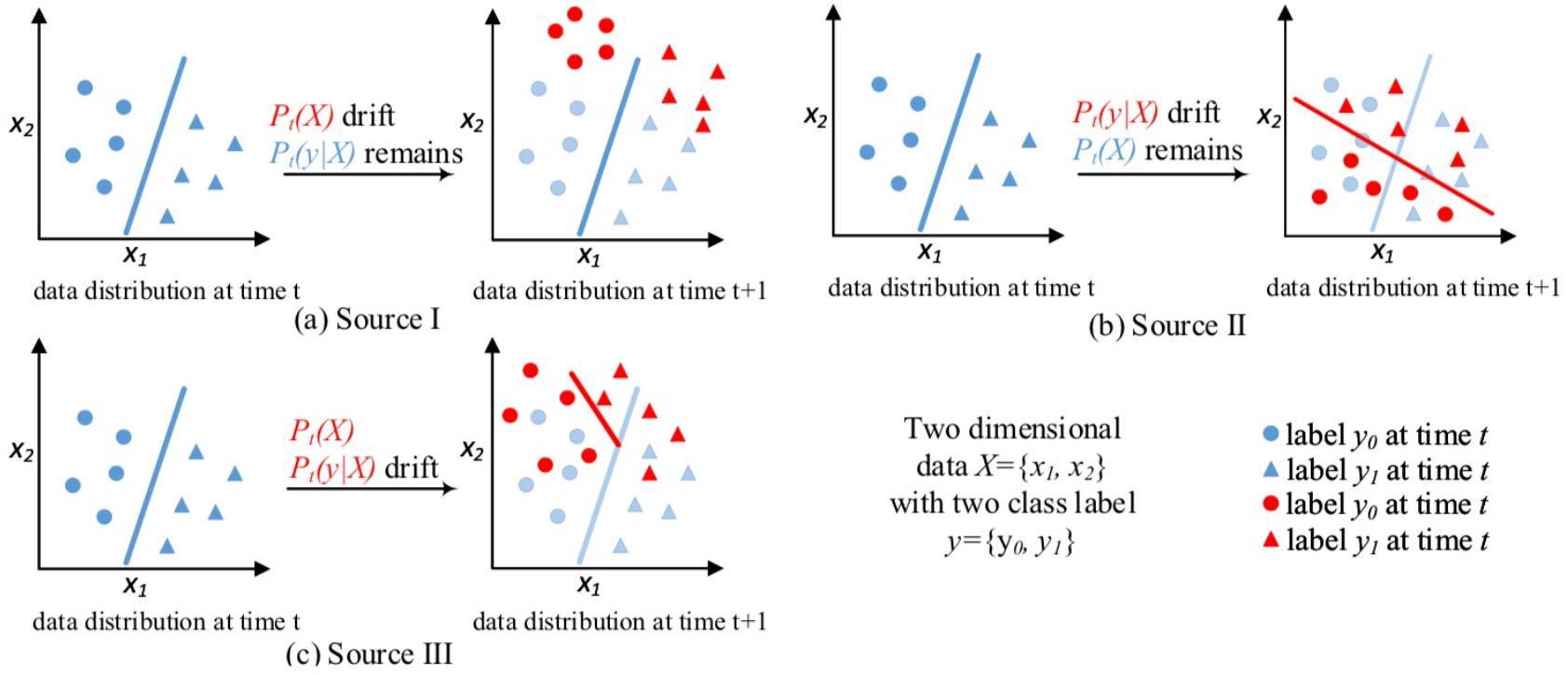


Fig. 15. A decision tree node is replaced with a new one as its performance deteriorates when a concept drift occurs in a subregion.

Before we start

Before we start



Before we start

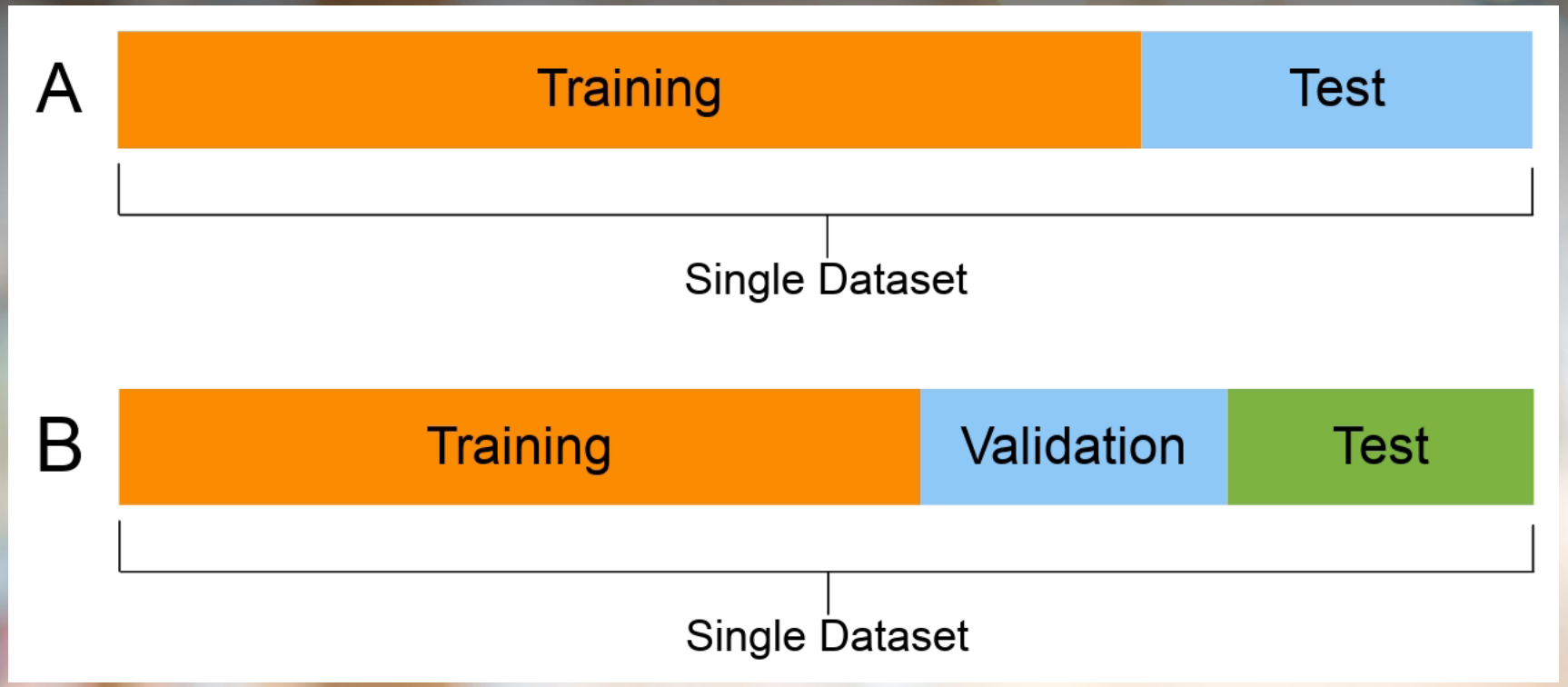
CONCEPT DRIFT

DATA DRIFT

Before we start

- Distribution เดิม concept เดิม
- Distribution เดิม concept ใหม่ (concept drift)
- Distribution ใหม่ concept เดิม (virtual drift)
- Distribution ใหม่ concept ใหม่ (data drift)

Before we start



Before we start

- ข้อมูลใน training set จะต้องประกอบด้วย distribution เดียวเท่านั้น
- ข้อมูลใน training set จะต้องประกอบด้วย concept เดียวเท่านั้น
- เพื่อทดสอบความ general, ข้อมูลใน training set และ test set ต้องอยู่ใน concept เดียวกัน
- ข้อมูลที่นำ model ไปใช้จริง จะต้องอยู่ใน concept เดียวกันกับ training set และ test set

Before we start

- ข้อมูลใน training set จะต้องประกอบด้วย distribution เดียวเท่านั้น
- ข้อมูลใน training set จะต้องประกอบด้วย concept เดียวเท่านั้น
- เพื่อทดสอบความ general, ข้อมูลใน training set และ test set ต้องอยู่ใน concept เดียวกัน
- ข้อมูลที่นำ model ไปใช้จริง จะต้องอยู่ใน concept เดียวกันกับ training set และ test set

Test 2 distributions are the same

Kolmogorov smirnov test	QQ plot	Runs test	Autocorrelation function
Shapiro-wilk test	Bartlett's test	Nist randomness tests	Anderson-darling test

Test 2 distributions are the same

$$P(X_1, X_2, \dots, X_p) = P(X_1 | X_2, \dots, X_p) P(X_2, X_3, \dots, X_p)$$

ให้ X ทุก feature เป็นอิสระต่อกัน

จะได้ว่า

$$P(X_1, X_2, \dots, X_p) = P(X_1) P(X_2, X_3, \dots, X_p)$$

จากสมบัติของ chain rule จะได้ว่า

$$P(X_1, X_2, \dots, X_p) = P(X_1) P(X_2) \dots P(X_p)$$

Test 2 concepts are the same

$$P(X_1, X_2, \dots, X_p, Y) = P(Y|X_1, X_2, \dots, X_p)P(X_1, X_2, X_3, \dots, X_p)$$

ให้ X ทุก feature เป็นอิสระต่อกัน

จะได้ว่า

$$P(X_1, X_2, \dots, X_p, Y) = P(Y|X_1, X_2, \dots, X_p)P(X_1)P(X_2) \dots P(X_p)$$

ถ้า X ของ 2 ชุดข้อมูลอยู่ใน distribution เดียวกัน จะได้ว่า

$$P(X_1, X_2, \dots, X_p, Y) = P(Y|X_1, X_2, \dots, X_p)$$

Test 2 concepts are the same

ในทางปฏิบัติพิจารณา

$$P(X_1, Y), P(X_2, Y), \ldots, P(X_p, Y)$$

เพื่อการทำความเข้าใจ concept ของข้อมูล


THANK YOU