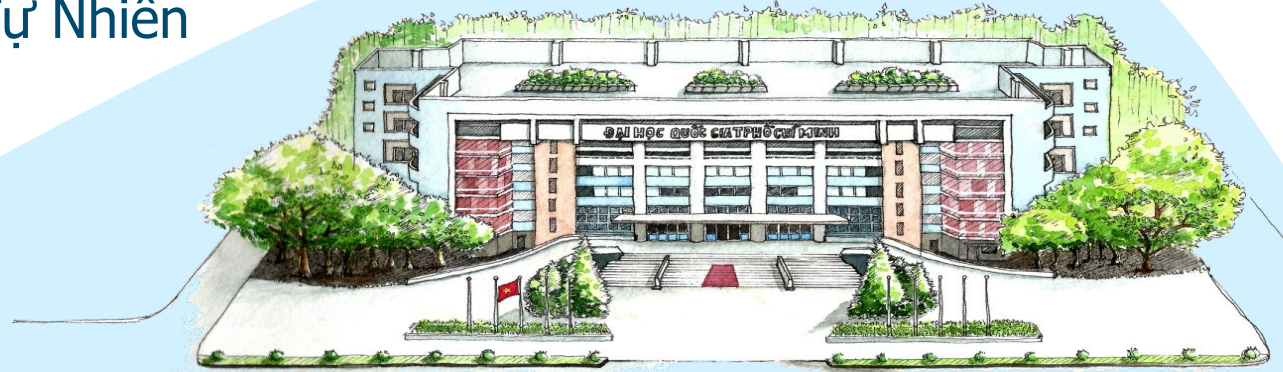


## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM



# Nội dung

**01**  
**Triển khai**

**02**  
**Giám sát**

**03**  
**Bảo trì**

# 01

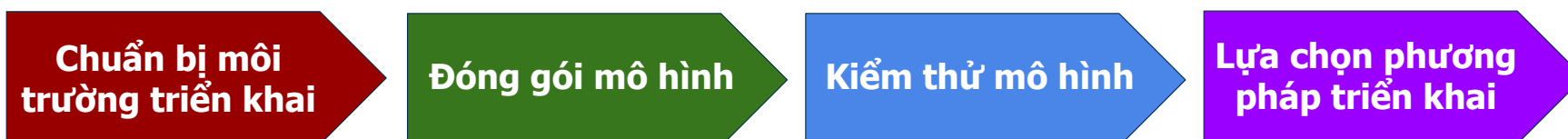
## Triển khai mô hình

1.1 Giới thiệu

1.2 Các bước chính để triển khai mô hình

## 1.1 Triển khai mô hình

- Là quá trình đưa mô hình đã phát triển vào môi trường thực tế để sử dụng
- Gồm các bước chính sau:



# 1.1 Triển khai mô hình

## Chuẩn bị môi trường triển khai

### Xác định nền tảng

**Đám mây(Cloud):** AWS, Google Cloud, Azure là những nền tảng phổ biến, cung cấp tài nguyên linh hoạt và dễ mở rộng.

**Máy chủ cục bộ (On-premise):** Phù hợp cho các tổ chức có dữ liệu nhạy cảm hoặc yêu cầu độ bảo mật cao.

**Edge Computing:** giúp giảm độ trễ, tăng tốc độ xử lý và tối ưu hóa băng thông mạng, phù hợp cho thiết bị IoT, thiết bị thông minh, xe tự lái.

### Kiểm tra tính tương thích

Đảm bảo môi trường triển khai tương thích với các yêu cầu của mô hình về phần mềm, framework, và tài nguyên tính toán.

### Thiết lập môi trường

Cài đặt các thư viện, phần mềm và tài nguyên (CPU, GPU) cần thiết cho mô hình hoạt động hiệu quả.

### Đảm bảo an ninh

Đảm bảo rằng các biện pháp bảo mật (quyền truy cập và mã hóa) đã được cấu hình đầy đủ.

## 1.1 Triển khai mô hình

### Đóng gói mô hình

#### Chuẩn bị container hoặc gói phần mềm (Sử dụng Docker)

- Đóng gói mô hình
  - Các thư viện cần thiết
  - Các tệp cấu hình
- vào một container → triển khai ở nhiều môi trường khác nhau.

#### Quản lý phiên bản (Versioning)

- Gắn nhãn phiên bản cho:
- Mô hình
  - Các tài liệu liên quan
- cập nhật khi cần thiết

## 1.1 Triển khai mô hình

### Kiểm thử mô hình

#### Kiểm tra tích hợp (Integration Testing)

Khả năng tương tác của mô hình với các hệ thống khác trong môi trường thực tế

#### Đánh giá hiệu suất

- Độ chính xác
  - Tốc độ xử lý
- đảm bảo các kết quả đạt yêu cầu.

## 1.1 Triển khai mô hình

### Lựa chọn phương pháp triển khai

#### Triển khai trực tiếp (Direct Deployment)

Phù hợp khi mô hình đã được tối ưu hóa và có hiệu suất ổn định.

#### Triển khai thử nghiệm (Shadow Deployment)

Chạy mô hình mới song song với mô hình hiện tại để so sánh hiệu suất trước khi triển khai chính thức.

#### Triển khai theo giai đoạn (Canary Deployment)

Triển khai mô hình cho một phần nhỏ người dùng trước khi mở rộng ra toàn bộ hệ thống, giúp giảm thiểu rủi ro khi thay đổi.



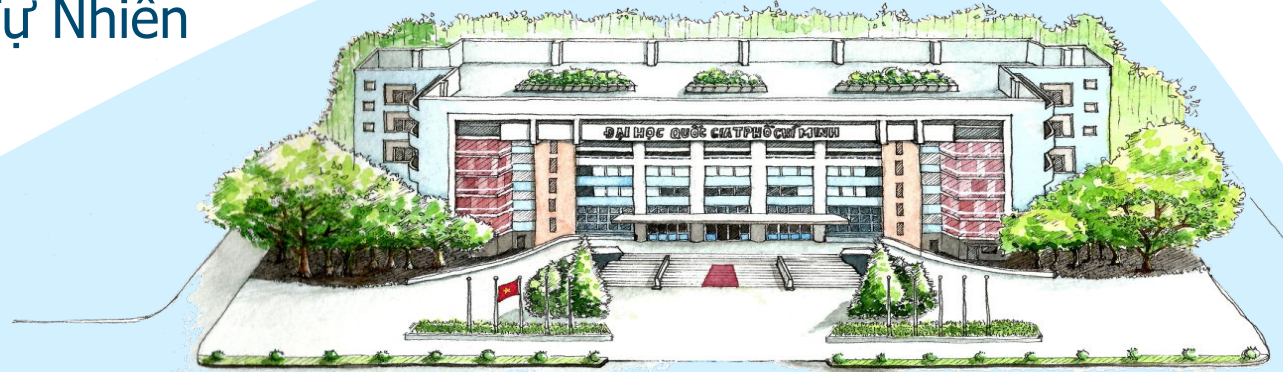
# Take-home message

- Triển khai mô hình và các điểm cần chú ý



## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM



# Nội dung



## Bài học trước

- Triển khai mô hình và các điểm quan trọng



## Bài học tiếp theo

- Giám sát
- Hiệu suất mô hình
- Độ trễ và thời gian phản hồi



# 02

## Giám sát

- 2.1 Hiệu suất mô hình
- 2.2 Độ trễ và thời gian phản hồi
- 2.3 Hiện tượng Model Drift
- 2.4 Giám sát sự cố
- 2.5 Thiết lập cảnh báo

## 2.1 Hiệu suất mô hình

- Theo dõi các chỉ số chính như
    - Độ chính xác
    - Độ nhạy (recall),
    - F1 score , ...
- đảm bảo mô hình duy trì chất lượng dự đoán.
- So sánh các chỉ số này với kết quả trong giai đoạn thử nghiệm.

## 2.2 Độ trễ và thời gian phản hồi

- Đánh giá tốc độ phản hồi của mô hình
  - quan trọng với ứng dụng yêu cầu thời gian thực.
- Nếu độ trễ tăng
  - → tối ưu hóa lại mô hình
  - → kiểm tra lại cơ sở hạ tầng

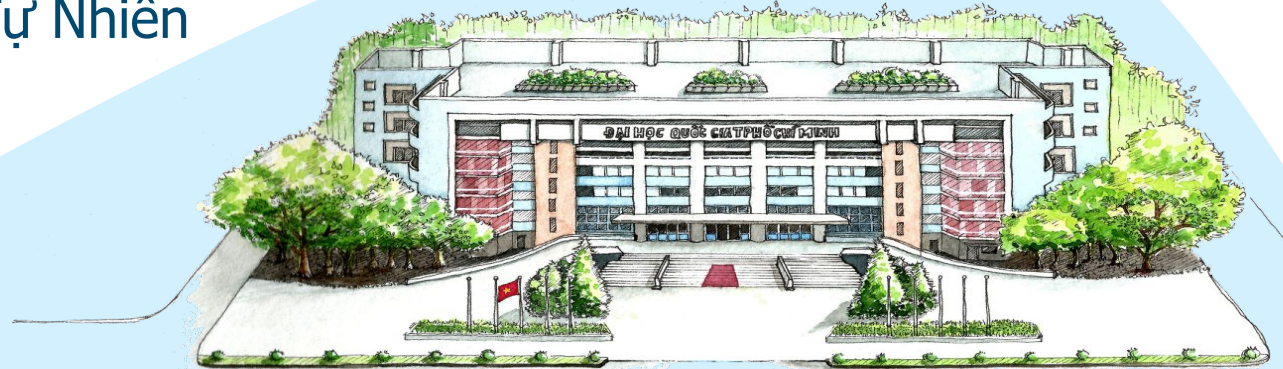
# Take-home message

- Giám sát
- Hiệu suất mô hình
- Độ trễ và thời gian phản hồi



## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM





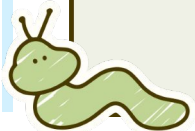
# Nội dung

## Bài học trước

- Giám sát
- Hiệu suất mô hình
- Độ trễ và thời gian phản hồi

## Bài học tiếp theo

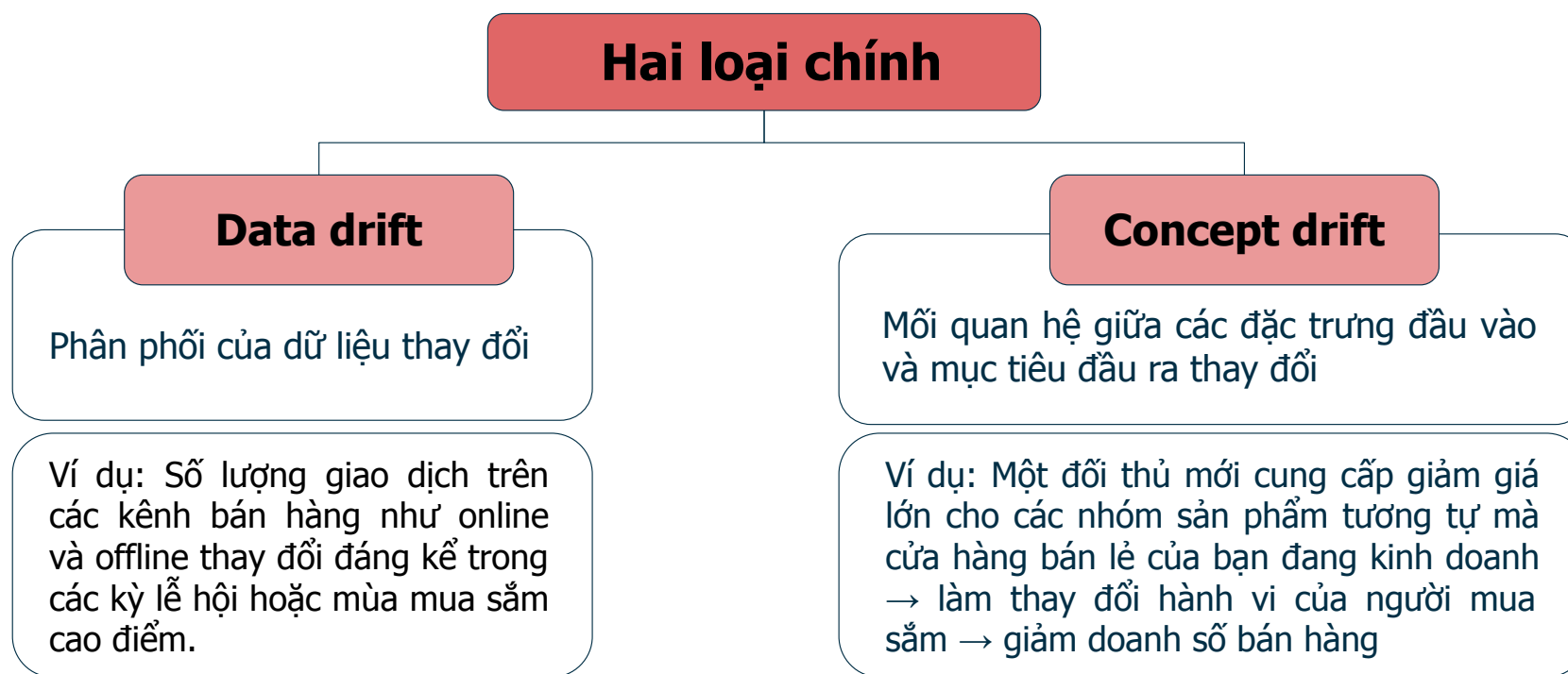
- Hiện tượng Model Drifting



## 2.3 Hiện tượng Model Drift

- Là hiện tượng xảy ra khi một mô hình máy học mất đi độ chính xác và khả năng dự đoán của nó theo thời gian.
  - Nguyên nhân:
    - Sự thay đổi trong dữ liệu đầu vào
    - Mối quan hệ giữa các biến trong dữ liệu thay đổi
- mô hình không còn phản ánh chính xác thực tế hiện tại như khi nó được huấn luyện ban đầu.

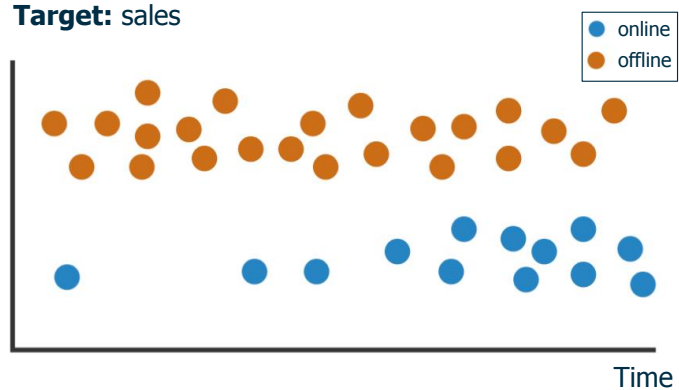
## 2.3 Hiện tượng Model Drift



## 2.3 Hiện tượng Model Drift

### Data drift

Target: sales

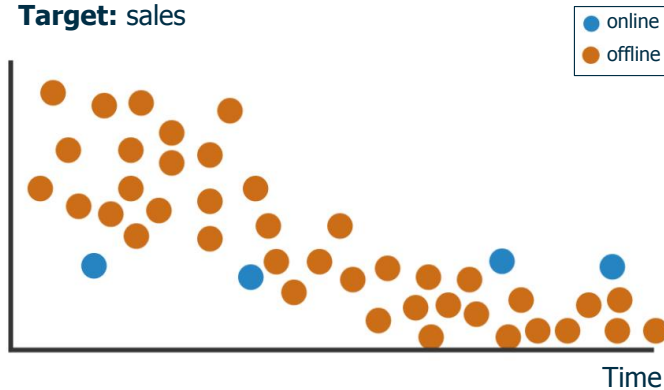


Feature distribution: sales channel



### Concept drift

Target: sales



Feature distribution: sales channel



## 2.3 Hiện tượng Model Drift

### Các cách khắc phục:

- Tái huấn luyện mô hình thường xuyên
- Sử dụng các thuật toán thích ứng
- Thiết kế lại mô hình

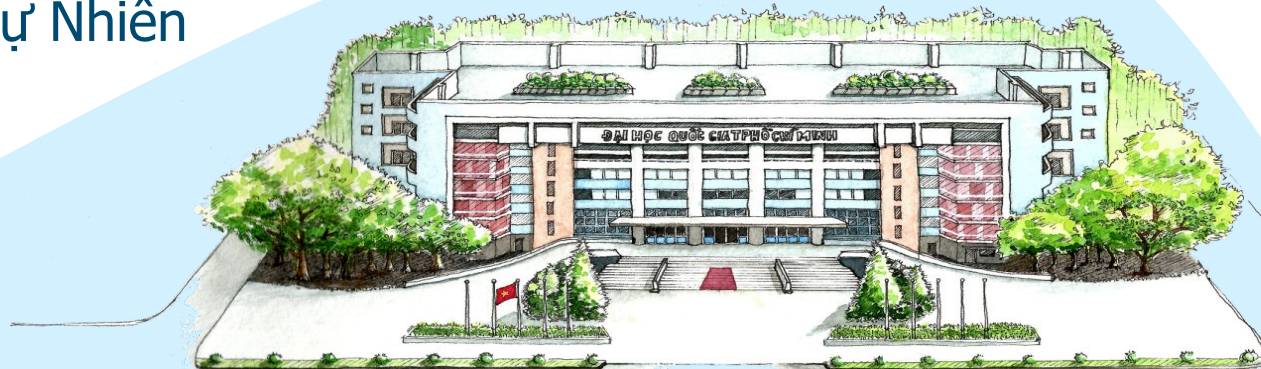
# Take-home message

- Model Drifting và các phương pháp xử lý



## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM



# Nội dung



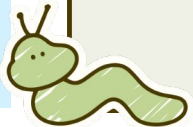
## Bài học trước

- Hiện tượng Model Drifting



## Bài học tiếp theo

- Giám sát sự cố
- Thiết lập cảnh báo





## 2.4 Giám sát sự cố

### Phát hiện và ghi lại:

- Các trường hợp sai lệch
  - Dự đoán sai bất thường
- quan trọng với các mô hình yêu cầu độ chính xác cao.

## 2.5 Thiết lập cảnh báo

- Đặt các ngưỡng cho từng chỉ số giám sát
- Thiết lập cảnh báo tự động khi chỉ số vượt quá giới hạn cho phép
- Sử dụng các công cụ giám sát phổ biến như Prometheus, Grafana hoặc các nền tảng đám mây như AWS CloudWatch, Google Cloud Monitoring để thiết lập cảnh báo tự động.

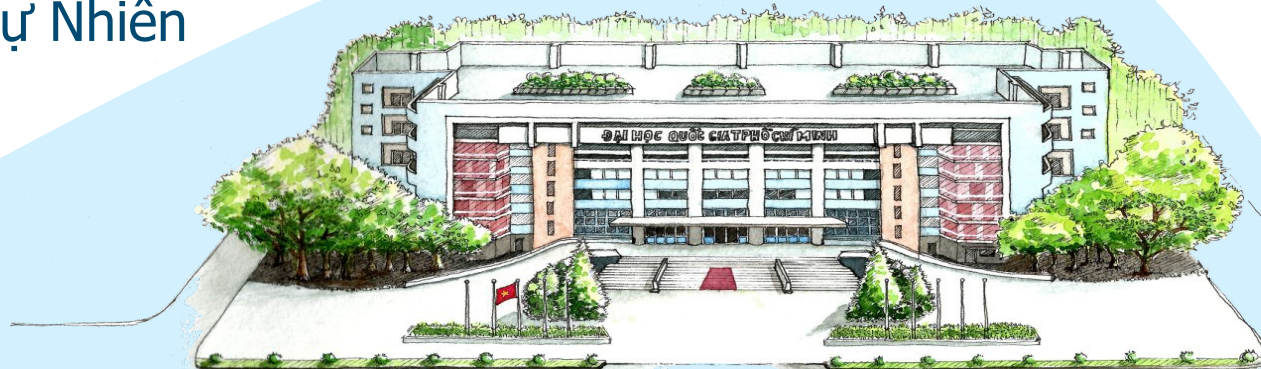
# Take-home message

- Giám sát sự cố
- Thiết lập cảnh báo



## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM



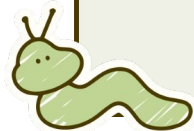
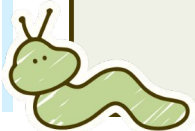
# Nội dung

## Bài học trước

- Giám sát sự cố
- Thiết lập cảnh báo

## Bài học tiếp theo

- Bảo trì mô hình
- Tái huấn luyện mô hình



# 03

## Bảo trì mô hình

- 3.1 Tái huấn luyện định kỳ
- 3.2 Cập nhật dữ liệu huấn luyện
- 3.3 Tối ưu hóa mô hình

## 3.1 Tái huấn luyện mô hình

- Tái huấn luyện:
  - Định kỳ với dữ liệu mới
  - Khi xảy ra hiện tượng Model Drift

→ Để cập nhật và duy trì độ chính xác.

Đối với các mô hình nhạy cảm với thời gian, tái huấn luyện cần thực hiện thường xuyên hơn.

## 3.2 Cập nhật dữ liệu huấn luyện

- Khi có dữ liệu mới
  - Dữ liệu đầu vào thay đổi
- Cập nhật tập huấn luyện giúp cải thiện hiệu suất mô hình
- Đảm bảo rằng tập dữ liệu vẫn phản ánh chính xác các xu hướng hiện tại.



## 3.3 Tối ưu hóa mô hình

- Dựa vào kết quả giám sát, có thể:
    - Tối ưu các siêu tham số
    - Cải tiến cấu trúc
    - Tối ưu mã nguồn
- giúp mô hình chạy nhanh, hiệu quả hơn.

# Take-home message

- Các phương pháp để cập nhật dữ liệu và tối ưu hoá mô hình



## 4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

PGS. TS. Nguyễn Thanh Bình  
Trưởng Đại học Khoa Học Tự Nhiên  
Đại học Quốc Gia Tp.HCM



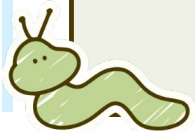
# Nội dung

## Bài học trước

- Cập nhật dữ liệu huấn luyện
- Tối ưu hoá mô hình

## Bài học tiếp theo

- Model Life Cycle



# Tóm tắt nội dung

## 01 Triển khai mô hình

## 02 Giám sát mô hình

- Đánh giá hiệu suất
- Model Drift
- Giám sát sự cố

## 03 Bảo trì mô hình

# Model Life Cycle

