

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

PGS. TS. Nguyễn Thanh Bình
Trường 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:

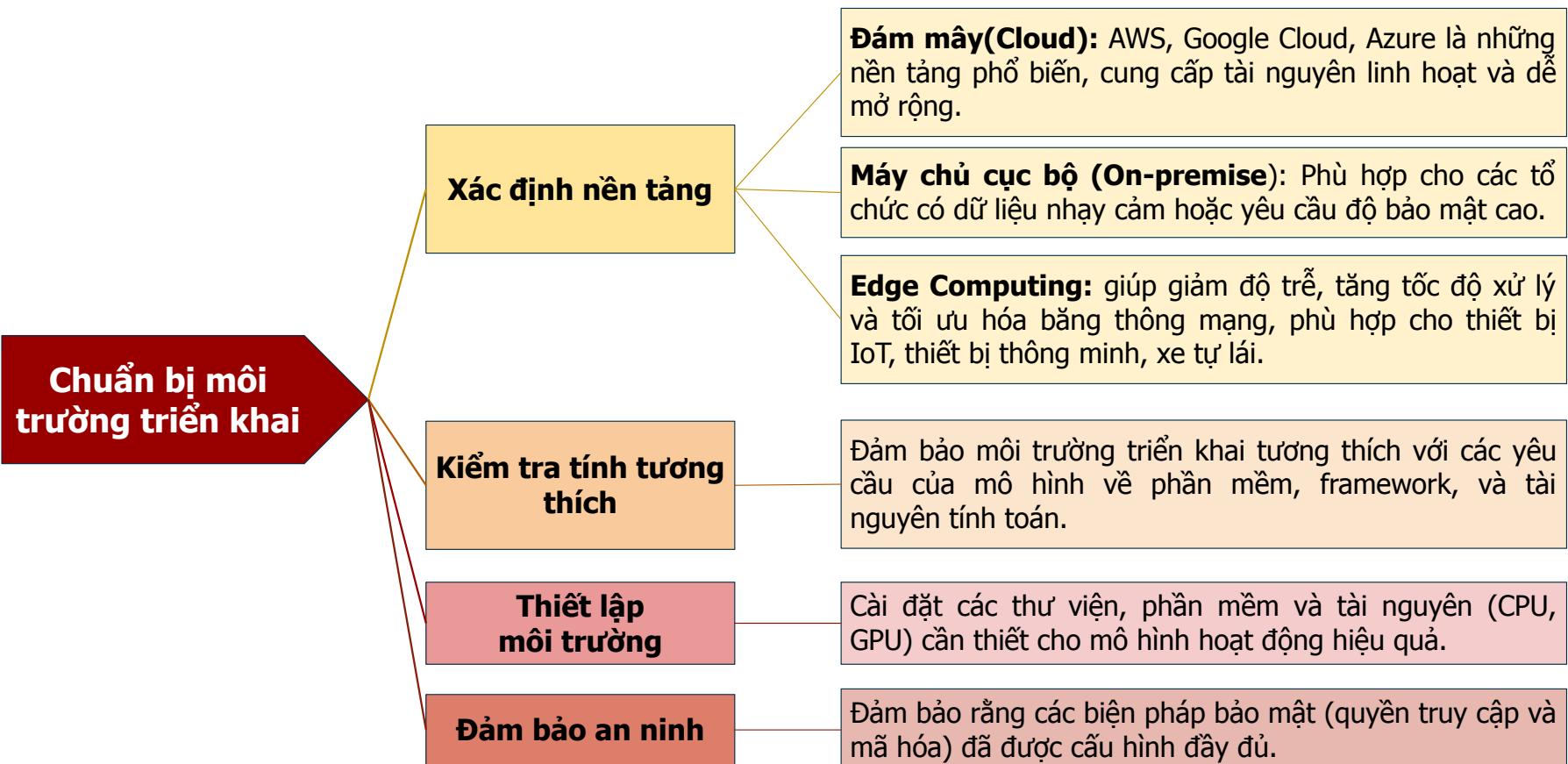
Chuẩn bị môi
trường triển khai

Đóng gói mô hình

Kiểm thử mô hình

Lựa chọn phương
pháp triển khai

1.1 Triển khai mô hình



1.1 Triển khai mô hình

Đóng gói mô hình

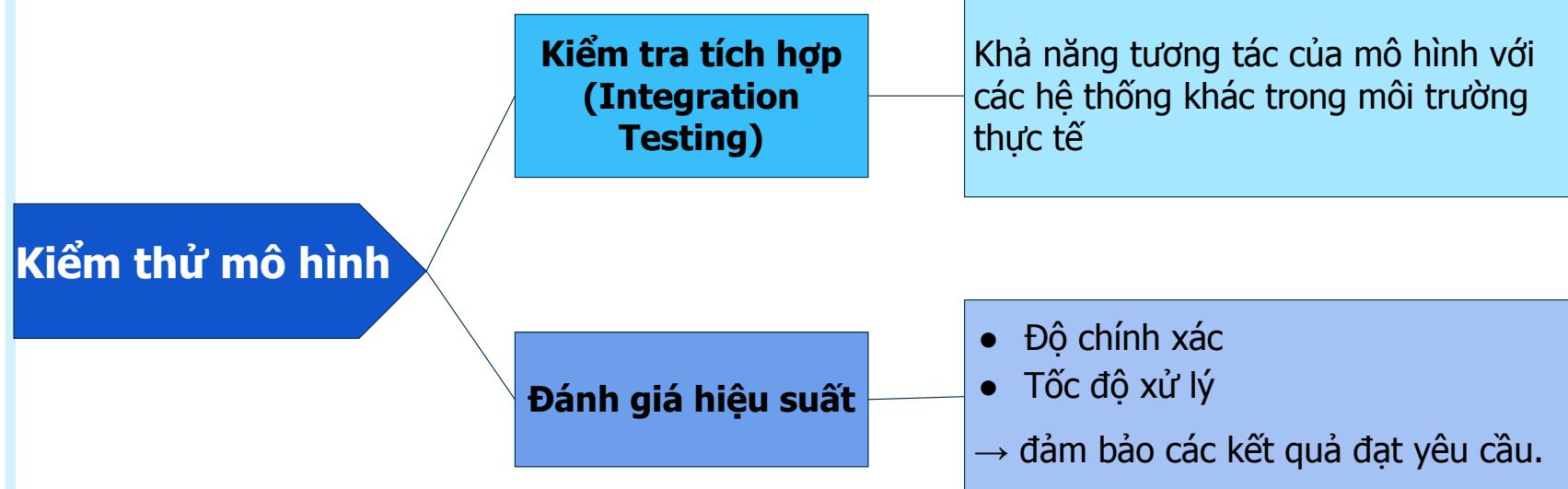
Chuẩn bị container
hoặc gói phần mềm
(Sử dụng Docker)

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

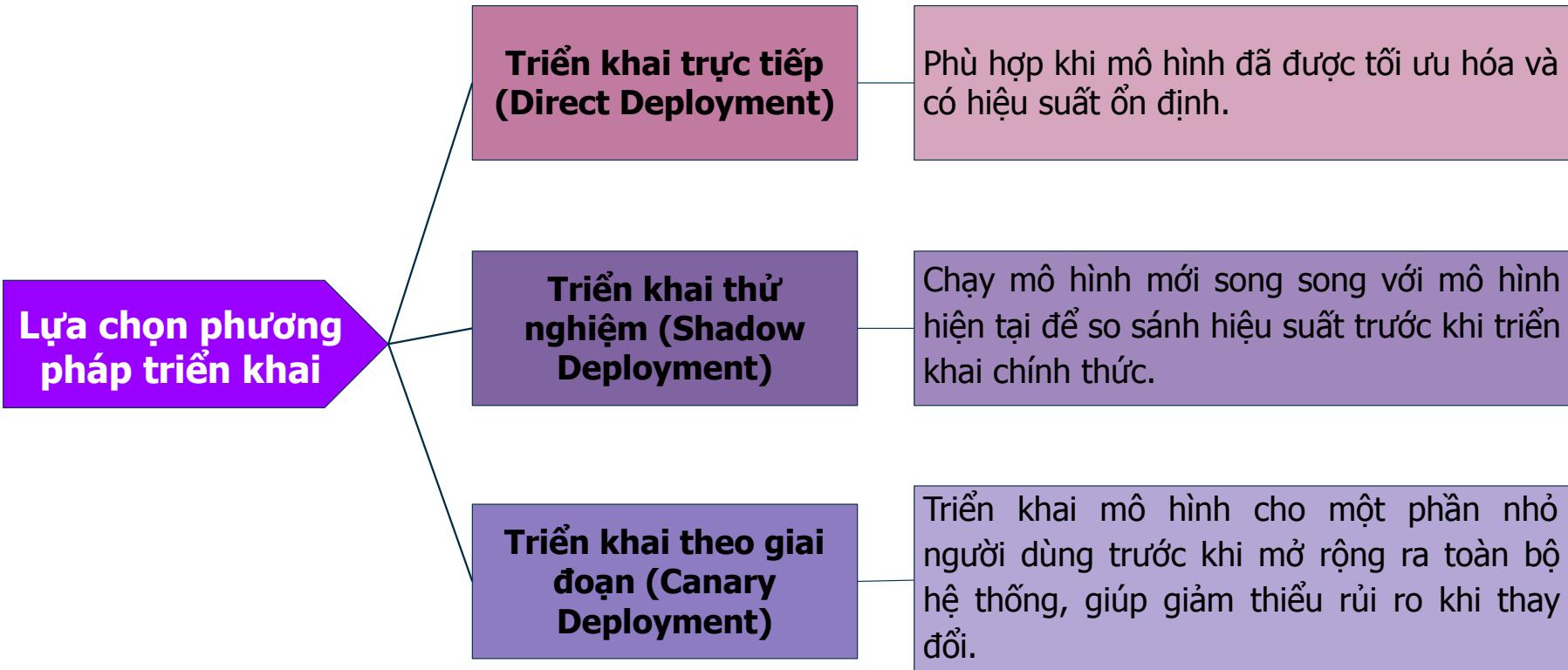
Quản lý phiên bản
(Versioning)

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

1.1 Triển khai mô hình



1.1 Triển khai mô hình



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
Trường Đại học Khoa Học Tự Nhiên
Đại học Quốc Gia Tp.HCM



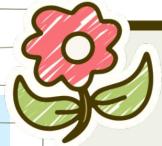
Nội dung

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
Trường 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

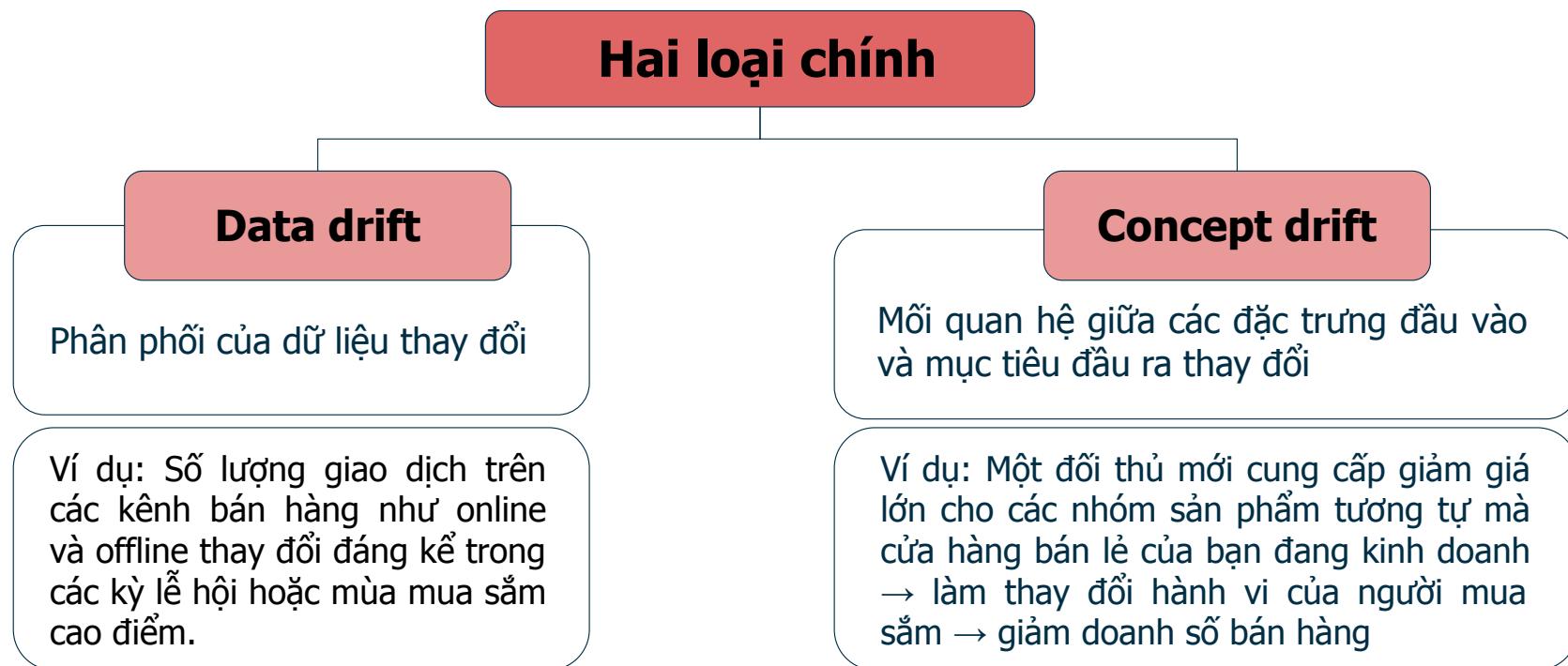
- Hiện tượng Model Drifting



2.3 Hiện tượng Model Drift

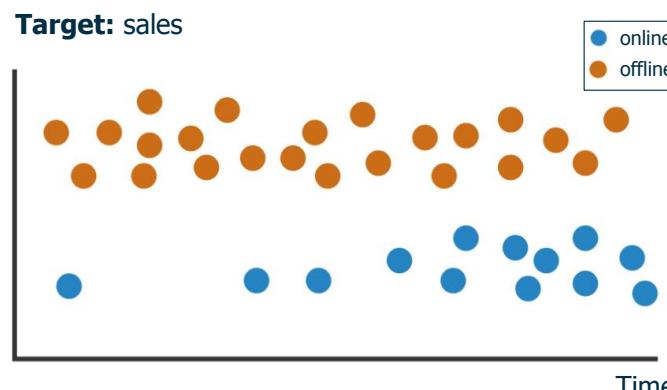
- Là hiện tượng xảy ra khi một mô hình máy học mất đi độ chính xác và khả năng dự đoán của nó theo thời gian.
- Nguyên nhân:
 - Sự thay đổi trong dữ liệu đầu vào
 - Mỗi quan hệ giữa các biến trong dữ liệu thay đổi→ mô hình không còn phản ánh chính xác thực tế hiện tại như khi nó được huấn luyện ban đầu.

2.3 Hiện tượng Model Drift



2.3 Hiện tượng Model Drift

Data drift

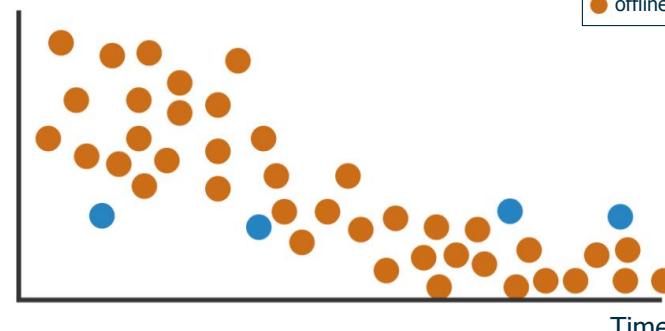


Feature distribution: sales channel



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
Trường Đại học Khoa Học Tự Nhiên
Đại học Quốc Gia Tp.HCM



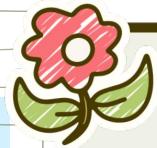
Nội dung

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

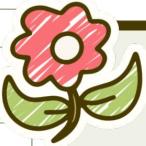


4.4 Triển khai, giám sát và bảo trì mô hình dữ liệu

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



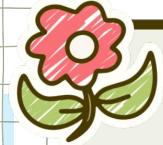
Nội dung

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

