1. **Machine learning là gì? So sánh nó với lập trình truyền thống**.

Machine learning một lĩnh vực của trí tuệ nhân tạo liên quan đến việc nghiên cứu và xây dựng các kĩ thuật cho phép các hệ thống "học" tự động từ dữ liệu để giải quyết những vấn đề cụ thể.

So sánh:

+ Lập trình truyền thống đòi hỏi phải tạo ra các hướng dẫn chi tiết để máy tính làm theo, việc viết chương trình để máy làm theo sẽ tốn nhiều thời gian hoặc không thể. Khi thay đổi kiểu bài toán ta phải viết thêm rule hoặc có khi phải viết lại toàn bộ chương trình.

+ Machine learning có cách tiếp cận là để máy tính học cách tự lập trình thông qua trải nghiệm. Khi thay đổi kiểu bài toán ta chỉ cần đưa thêm kiểu dữ liệu mới vào cho nó.

1. **Labeled data là gì? Nêu tên các bài toán thuộc loại Supervised learning và cho ví dụ cụ thể (khác với ví dụ trong bài giảng). Giải thích tại sao đó là Supervised learning.**

Labeled data là dữ liệu có nhãn mà labeled là câu trả lời của bài toán

Tên các bài toán thuộc loại Supervised learning:

+ Classiification: ví dụ bài toán nhận diện các khuôn mặt trong một bức ảnh như tính năng gắn thẻ trên fb vì số lượng cặp dữ liệu (khuôn mặt, tên người) là riêng biệt với nhau

+ Regression: ví dụ bài toán dự đoán tuổi dựa trên khuôn mặt nếu ta coi tuổi là một dãi số nguyên dương từ 1 đến 100

1. **Nêu tên các bài toán thuộc loại Unsupervised learning và cho ví dụ cụ thể (khác với ví dụ trong bài giảng). Giải thích tại sao đó là Unsupervised learning.**

Clustering: ví dụ đưa cho đứa trẻ nhiều mảnh ghép với các hình dáng và màu sắc khác nhau thì chúng vẫn có khả năng phân loại theo màu sắc và hình dáng dù không biết mảnh đó hình gì hay màu gì

Anomaly detection: ví dụ bài toán phát hiện gian lận là hệ thống sẽ kiểm tra số lần click chuột, số lần gửi request tới server, thời gian giữa các request…của người dùng, từ đó hệ thống có thể phát hiện người dùng có các hành động bất thường và đưa ra cảnh báo với người quản trị hệ thống

Dimensionality reduction: ví dụ là khi người ta mua quần áo thường sẽ mua thêm giày hoặc nón nên hệ thống sẽ gợi ý nhờ đó thúc đẩy nhu cầu mua sắm

1. **Semi-supervised learning là gì. Cho ví dụ về Semi-supervised learning.**

Semi-supervised là một nửa giám sát của con người một nửa không.

Ví dụ: kho lưu trữ hình ảnh thì một số được gán nhãn một số thì không

1. **Batch learning là gì? Online learning là gì? Cho ví dụ mỗi loại.**

Batch learning: hệ thống hoàn tất rồi mới đưa ra cho người khác sử dụng

Ví dụ: phần mềm dùng để học offline

Online learning: hệ thống có dữ liệu thay đổi liên tục nên vừa được sử dụng vừa cập nhật liên tục

Ví dụ: phần mềm học online sẽ được cập nhật dữ liệu liên tục

1. **C**

Huấn luyện robot dọn dẹp nhà cửa sử dụng Reinforcement learning là vì đầu tiên ta sẽ chỉ ra nơi cần dọn dẹp cho robot, sau đó robot sẽ tự học được và tự xác định được nơi cần dọn dẹp

1. **D**

Để nhận dạng virus máy tính ta có thể dùng:

+ Supervised learning vì ta có thể đưa ra trước rất nhiều loại virus cho hệ thống biết để có thể nhận diện khi gặp

+ Unsupervised learning vì hệ thống sẽ phát hiện hành động bất thường gây hại cho máy tính của một cái gì đó và phát ra tín hiệu dù nó chưa từng được biết đó là virus

+ Reinforcement learning vì trước tiên ta sẽ chỉ ra đâu là virus cho nó biết rồi sẽ cộng điểm thưởng cho nó nên về sau để tối ưu điểm thưởng nó sẽ tự tìm ra virus