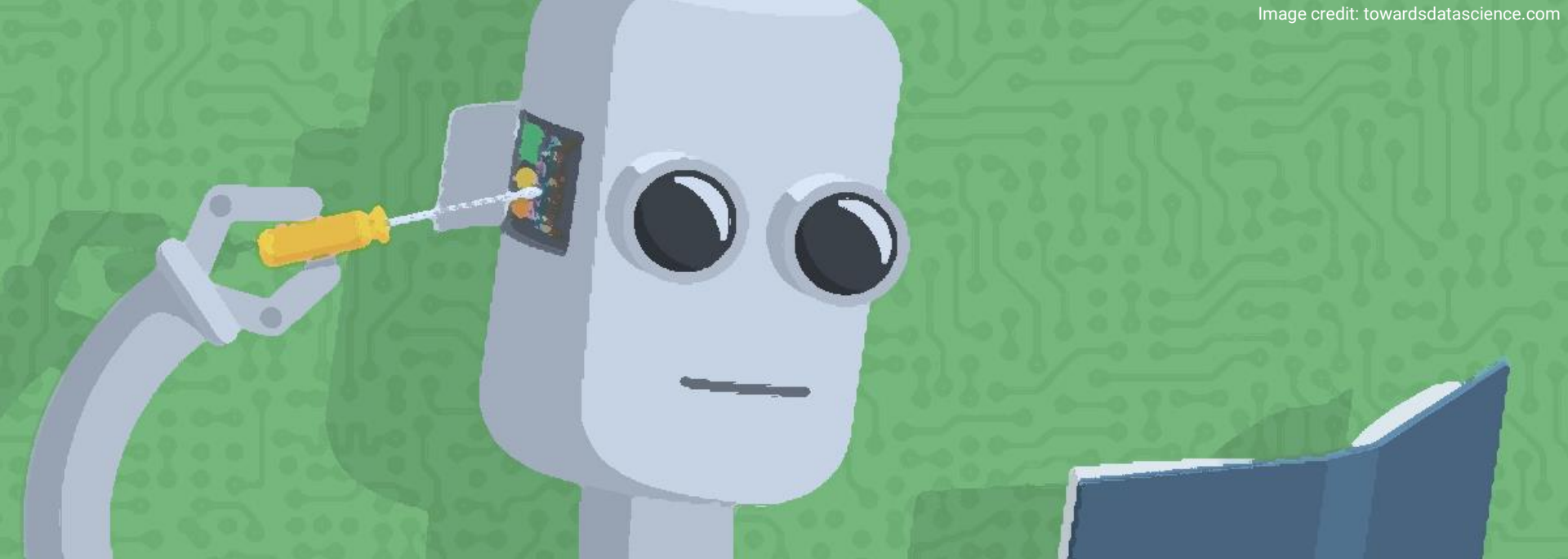


WEEK 7

Tìm hiểu về Machine learning



MACHINE LEARNING

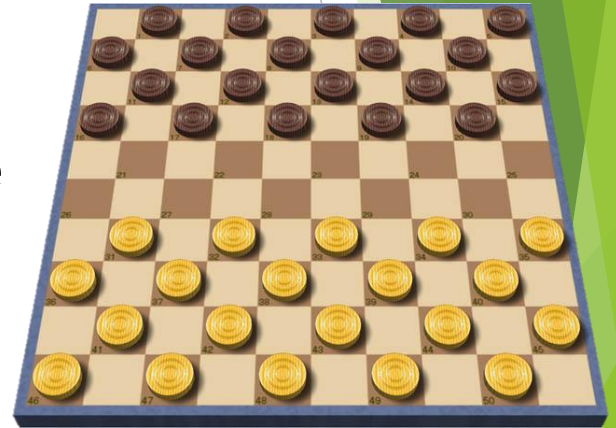


▶ WHAT IS MACHINE LEARNING?

Machine learning definition

Machine learning is a field of study that gives computers the ability to learn from experience **without being explicitly programmed.**

(Samuel, 1959)



Checkers

Image credit: Michel32NI

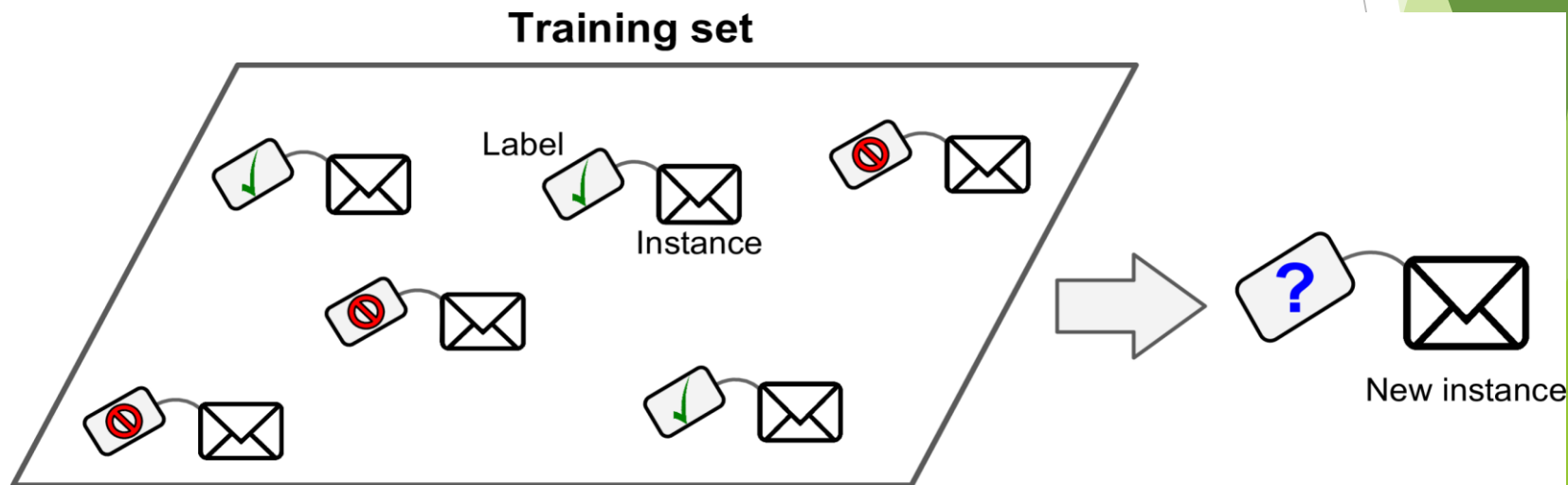
► TYPES OF ML

ML classification criteria

- ▶ 1. **With human supervision** or not?
 - ▶ supervised learning, unsupervised learning, semisupervised learning, and reinforcement learning
- ▶ 2. **Learn on the fly** or not?
 - ▶ online learning, batch learning
- ▶ 3. **Use model** or not?
 - ▶ instance-based learning, model-based learning

With human supervision?

- ▶ Supervised learning: use **labeled data**
- ▶ *Classification*, e.g., spam filter, image classification



Email classification



Image classification

airplane



automobile



bird



cat



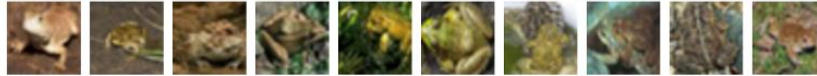
deer



dog



frog



horse



ship

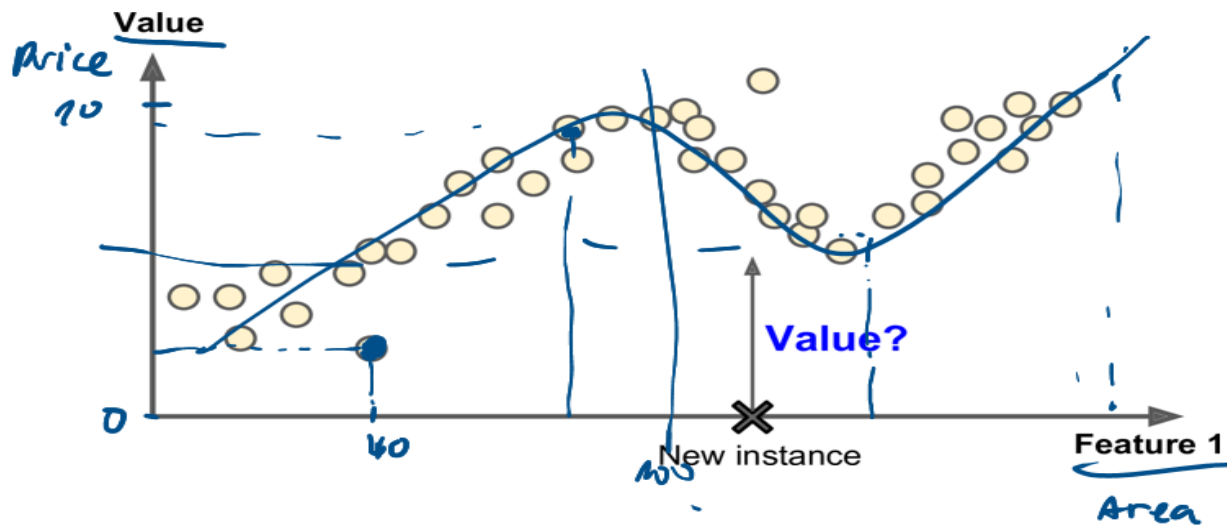


truck



With human supervision?

- Supervised learning: use **labeled data**
- *Regression*, e.g., housing price prediction, weather forecast



Weather forecast

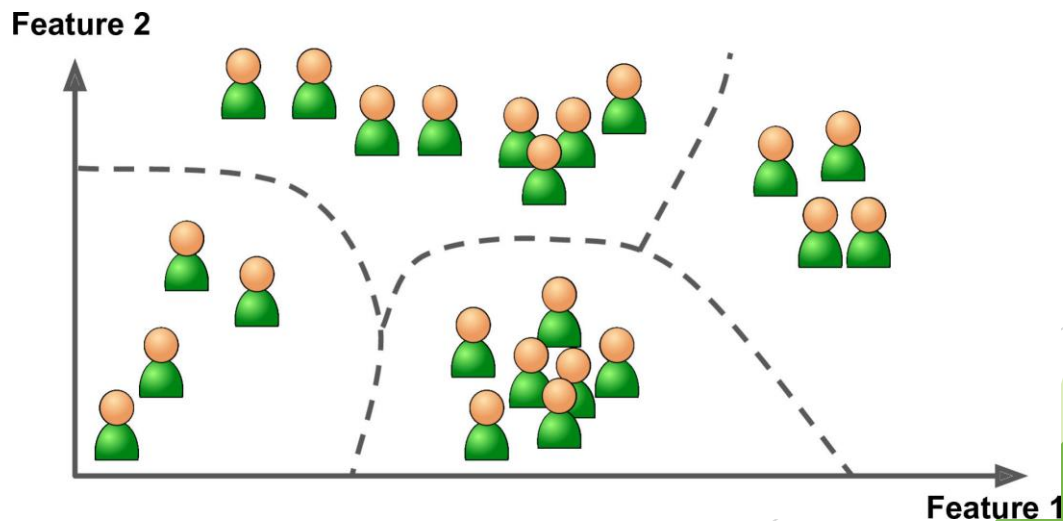


With human supervision?

- ▶ Supervised learning: most important algorithms include
 - ▶ k-Nearest Neighbors
 - ▶ Linear Regression
 - ▶ Logistic Regression
 - ▶ Support Vector Machines (SVMs)
 - ▶ Decision Trees and Random Forests
 - ▶ Neural networks

With human supervision?

- Unsupervised learning: use **unlabeled data**
 - *Clustering*: most important algorithms include K-Means, DBSCAN, Hierarchical Cluster Analysis (HCA)




News clustering

 Tin tức

Tìm kiếm chủ đề, vị trí và nguồn

 Tin bài hàng đầu

 Dành cho bạn

 Mục ưa thích

 Mục tìm kiếm đã lưu

 Việt Nam

 Thế giới

 Địa phương

 Doanh nghiệp

 Giải trí



Cuối cùng Windows 10 đã vượt Windows 7 trở thành hệ điều hành PC phổ biến nhất thế giới-Công nghệ thông tin

Tin tức 24h • hôm nay





- Cuối cùng Windows 10 cũng vượt Windows 7, trở thành hệ điều hành PC phổ biến nhất thế giới
GenK • hôm qua



- Thị phần Windows 10 lần đầu vượt Windows 7 | Sản phẩm - Dịch vụ
VietTimes • hôm nay



- Cuối cùng, số người dùng Windows 10 đã vượt qua Windows 7
VNReview • 5 giờ trước

- Windows 10 chính thức vượt mặt Windows 7 trở thành hệ điều hành phổ biến nhất
Người đưa tin • hôm nay

 Xem thông tin toàn cảnh



News clustering



Điểm thi ĐH 2020: tăng cao, xét tuyển nhiều rủi ro

Kênh Tuyển Sinh · 2 ngày trước



- Nhập học ít, thí sinh đang ở đâu?

Báo Thanh Niên · Hôm qua



- Điểm sàn của 15 trường, khoa đào tạo ngành sức khỏe

Zing.vn · Hôm qua



- Cập nhật 22/9: 80 trường đại học công bố điểm sàn, điểm chuẩn, thí sinh cần nắm rõ để thay đổi nguyện vọng kịp thời

Kênh 14 · Hôm qua

- Trúng tuyển 'ảo' quá nhiều, điểm chuẩn ĐH sẽ biến động thế nào?

VietNamNet · Hôm qua



Xem Thông tin toàn cảnh



Social network analysis



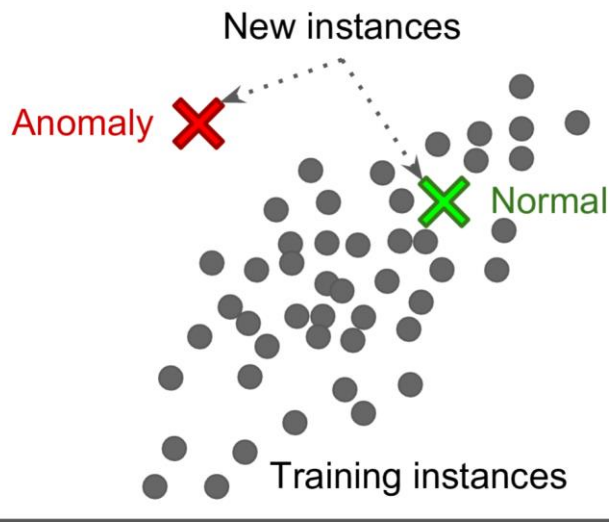
Market segmentation



With human supervision?

- Unsupervised learning: use **unlabeled data**
 - *Anomaly detection*: One-class SVM, Isolation Forest

Feature 2



Feature 1

Use model?

- Instance-based learning: **compares new data to the learned examples using a similarity measure.**

Feature 2



Use model?

- Model-based learning: **build a model** of the learned examples.

