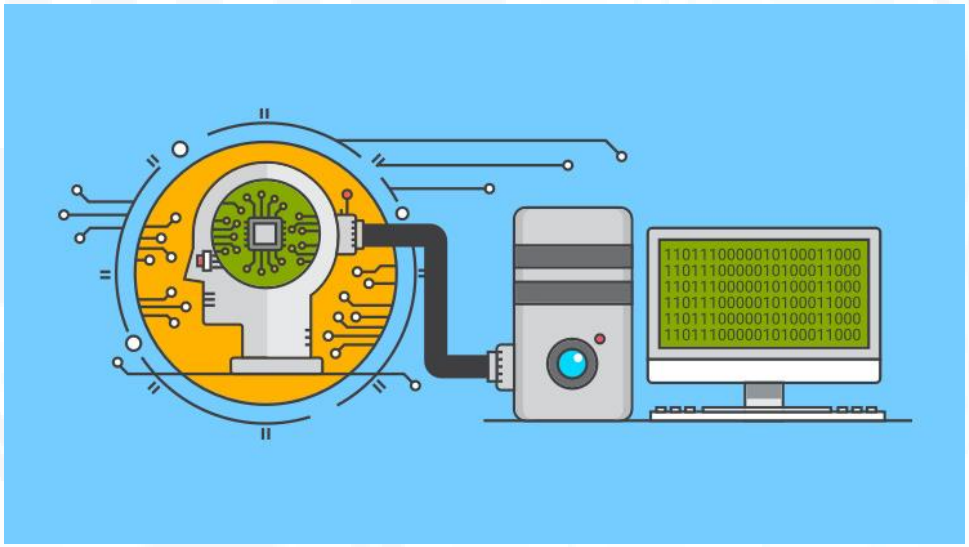
FACULTY OF INFORMATION TECHNOLOGY

Machine Learning

(Học Máy)

Semester 2, 2023/2024

Chapter 1. Introduction



Content

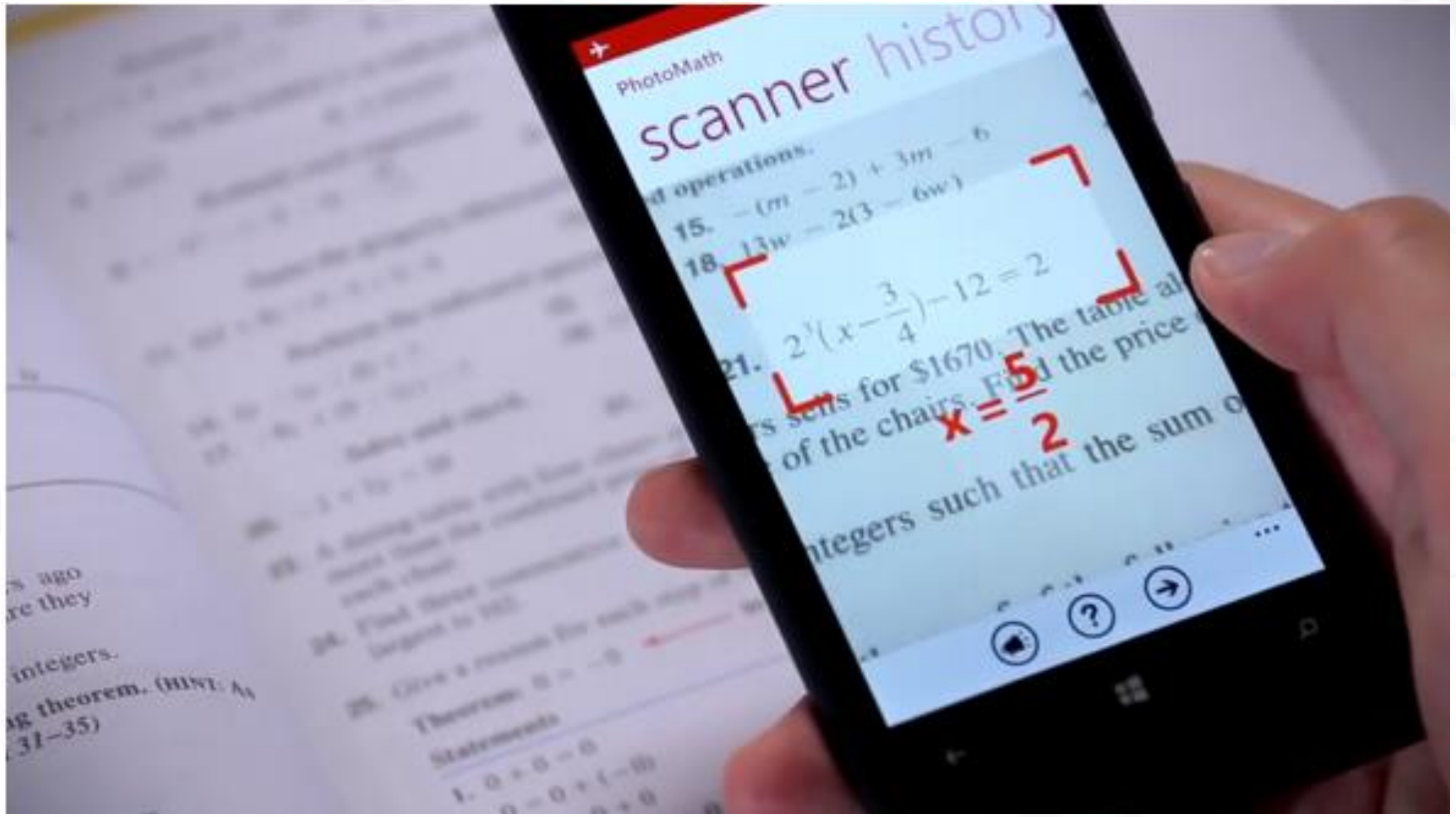
- What is Machine Learning?
- Machine Learning Types
- Machine Learning Algorithms
- Applications of Machine Learning

Situation 1

- How can we make a robot cook?

Situation 2

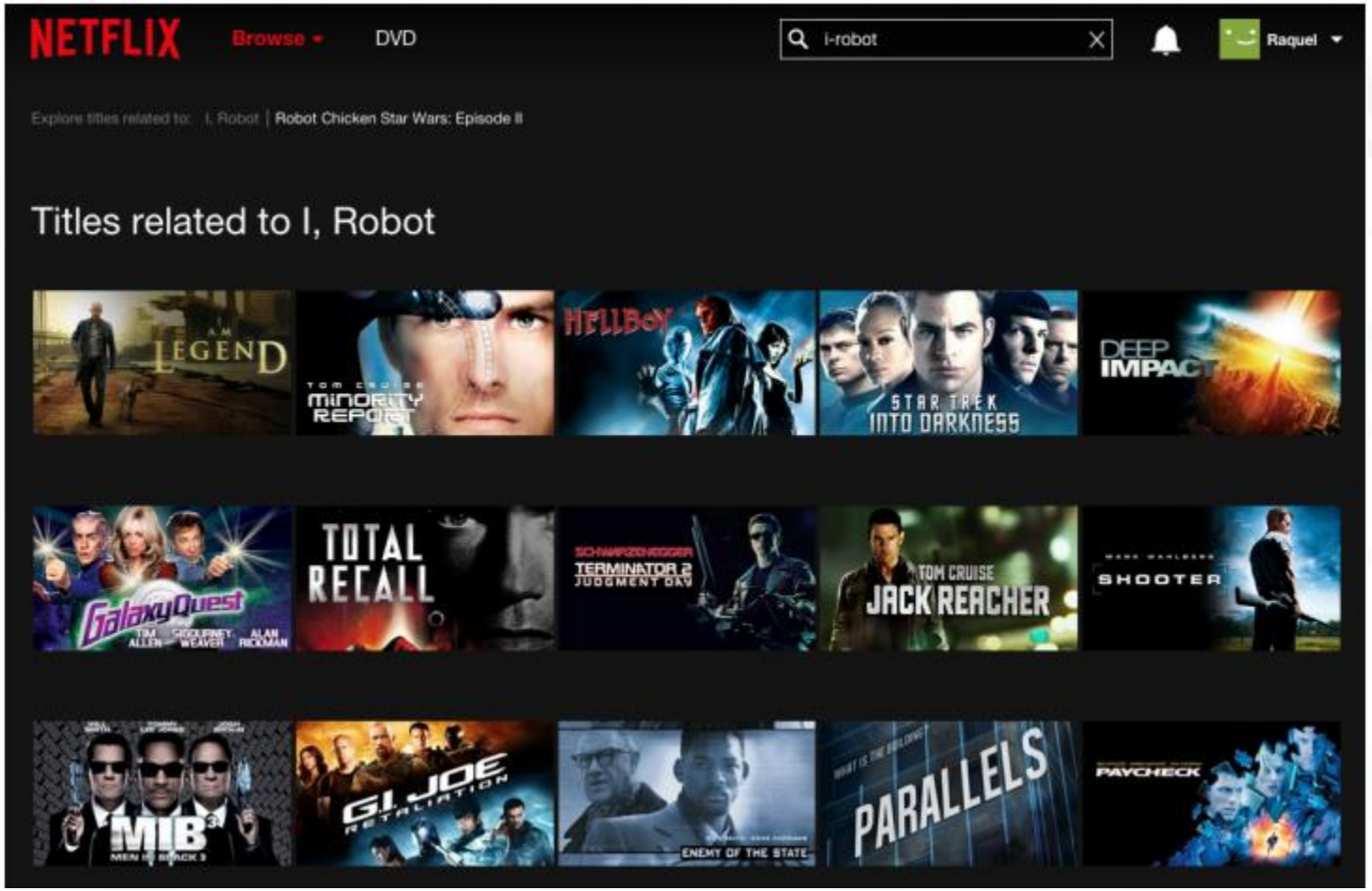
- How to recognize a math expression?



Photomath: https://photomath.net/

Situation 3

- How does Netflix recommend films to a user?



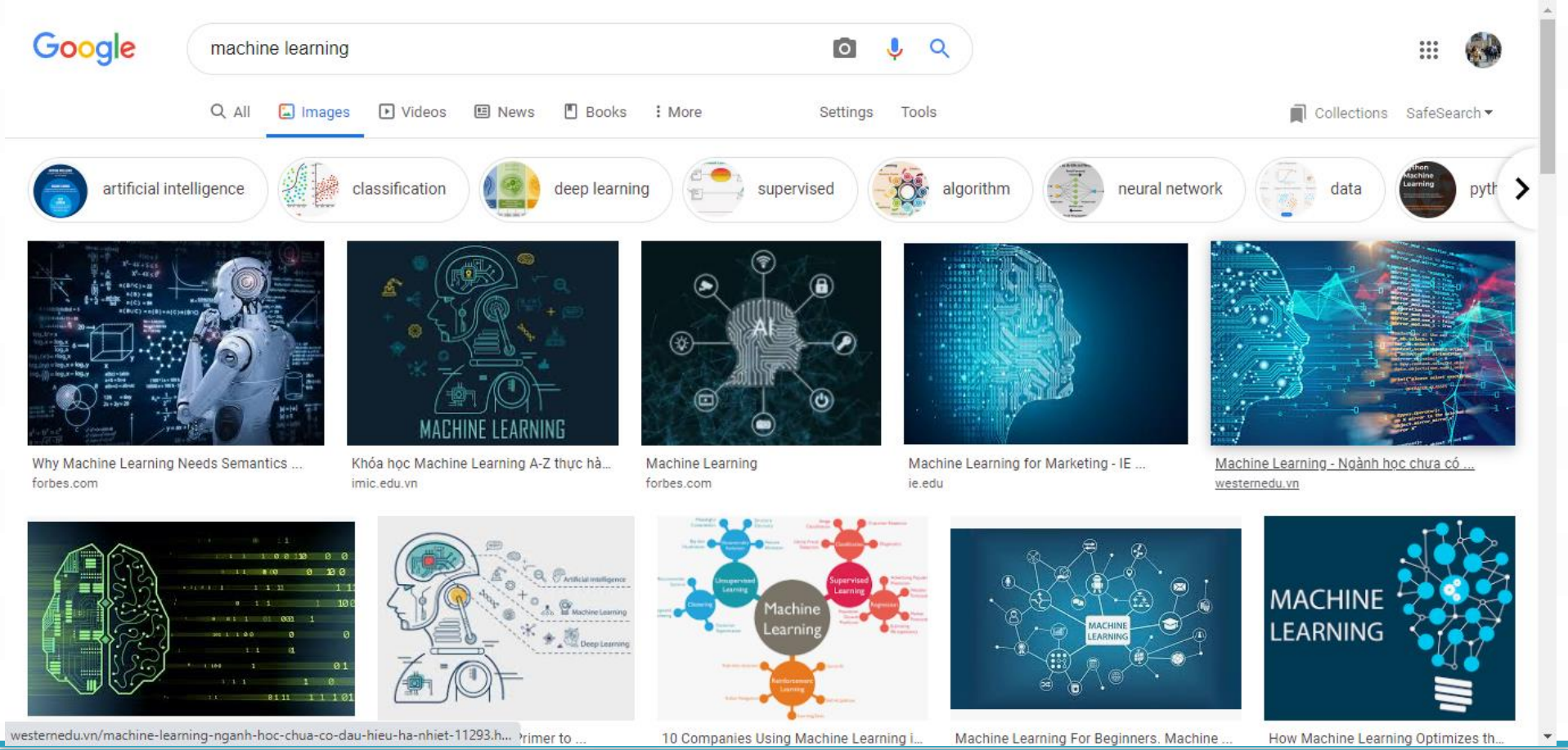
Situation 4

- How to predict the price of a house?

Situation 5

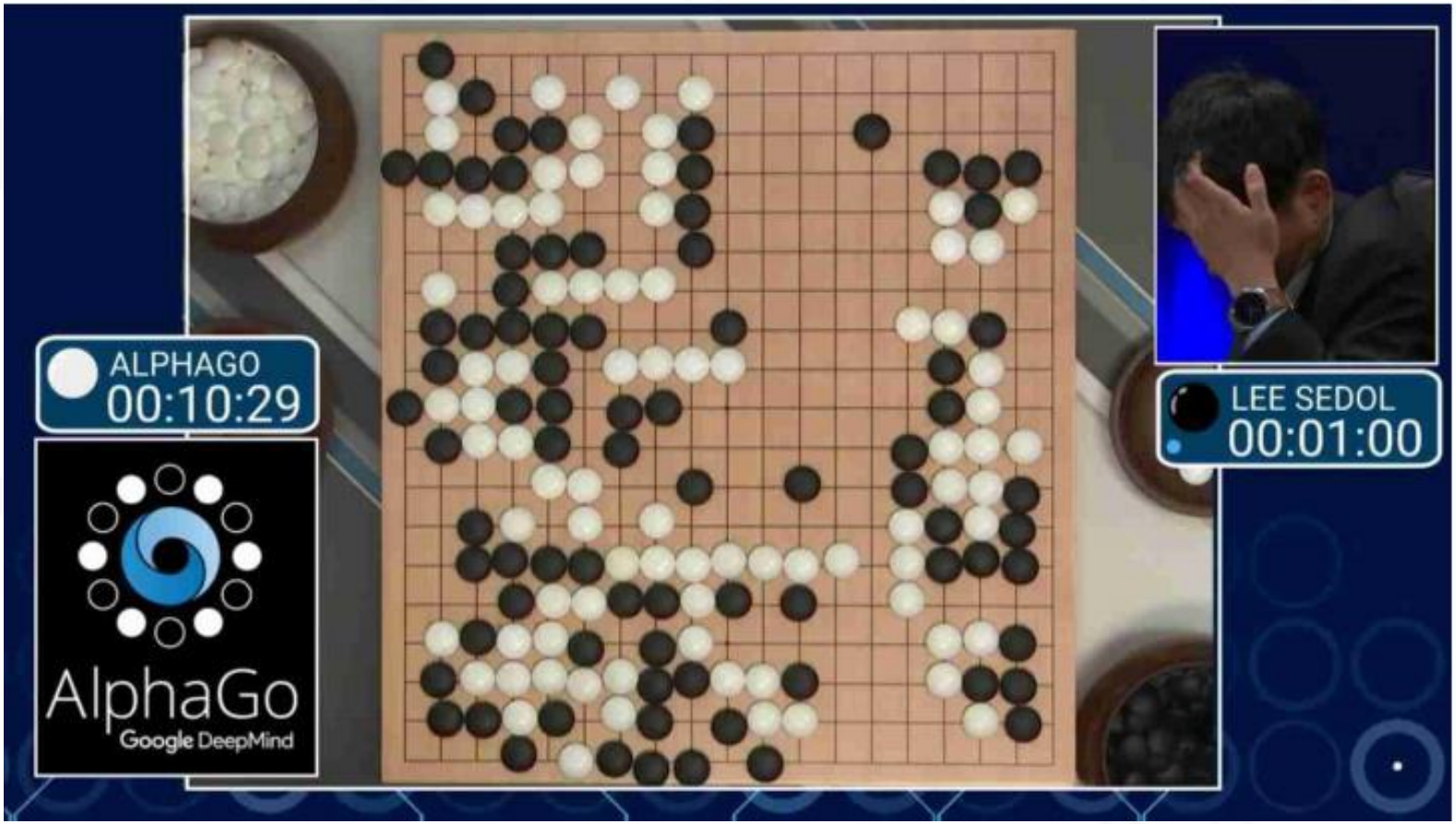
Information Retrieval

- How does Google search images based on a given text?



Situation 6

- How can AI play game?



Situation 7

- Am I going to pass the ML course?

Situation …

What is Learning?

- Dictionary defines "to learn" as:
  - To get knowledge of something by study, experience, or being taught.
  - To become aware by information or from observation
  - To commit to memory
  - To be informed of or to ascertain
  - To receive instruction

What is Learning? (cont.)

- A learning program: learn from experience E on task T with respect to performance measure P, if its performance on T improves with experience E.

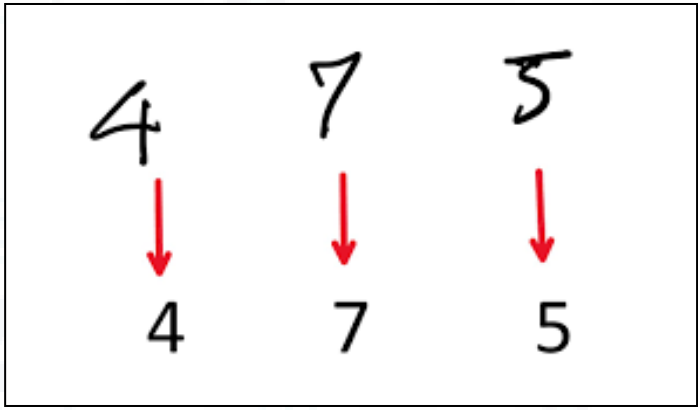
- A learning program:
  - produces a representation R (often called *a hypothesis h*) of what it has learned (or a **model**).
  - another program can use R to perform T.
  - uses a learning algorithm A to produce R from E.
  - many different algorithms can be used produce the same type of representation.

Example: Learning Checkers

- Task T:
  - playing checkers
- Performance measure P:
  - percent of games won against a set of players.
- Training experience E:
  - games played against the other players.
- Representation R:
  - an evaluation function that measures the goodness of the board.
- Learning algorithm A:
  - a variation of temporal difference (predicting the total reward expected over the future)

Example: Learning Handwriting

- Task T:
  - recognize handwritten letters.
- Performance measure P:
  - error rate on sample handwriting.
- Training experience E:
  - gray scale images of sample handwriting, all identified in advance.
- Representation R:
  - support vector machine, where each pixel is a separate attribute.
- Learning algorithm A:
  - SVMlight.



What is Machine Learning?

- Wikipedia: (ML introduced in 1980's)
  - *Machine learning is the subfield of computer science that "gives computers the ability to learn without being **explicitly programmed**"*

- Ability of computers to "learn" from "data" or "past experience"
  - learn: Make intelligent predictions or decisions based on data by optimizing a model
  - data: Comes from various sources such as sensors, domain knowledge, experimental runs, etc.

Why "Learning" ?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to "learn" to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

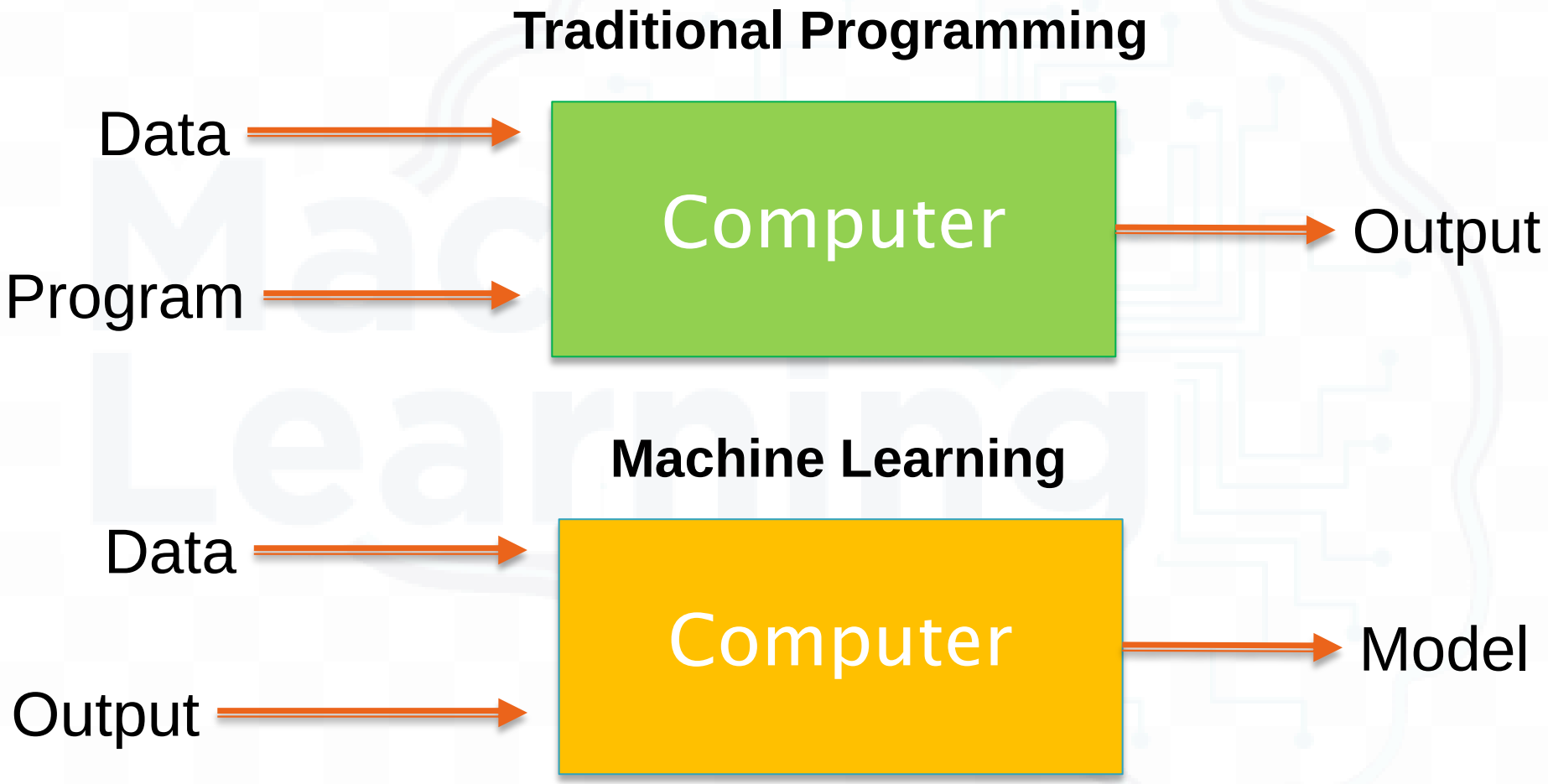
What We Talk About When We Talk About "Learning"?

- Learning general models from data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:

  People who bought "Blink" also bought "Outliers" (www.amazon.com)

- Build a model that is *a good and useful approximation* to the data.

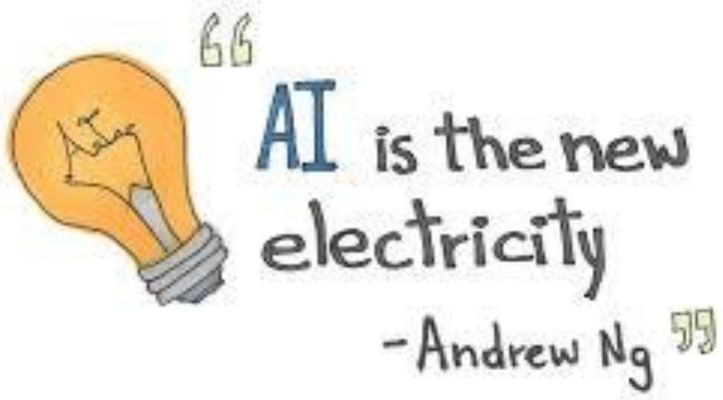
Machine Learning vs Traditional Programming



Artifical Intelligence vs Machine Learning vs Deep Learning

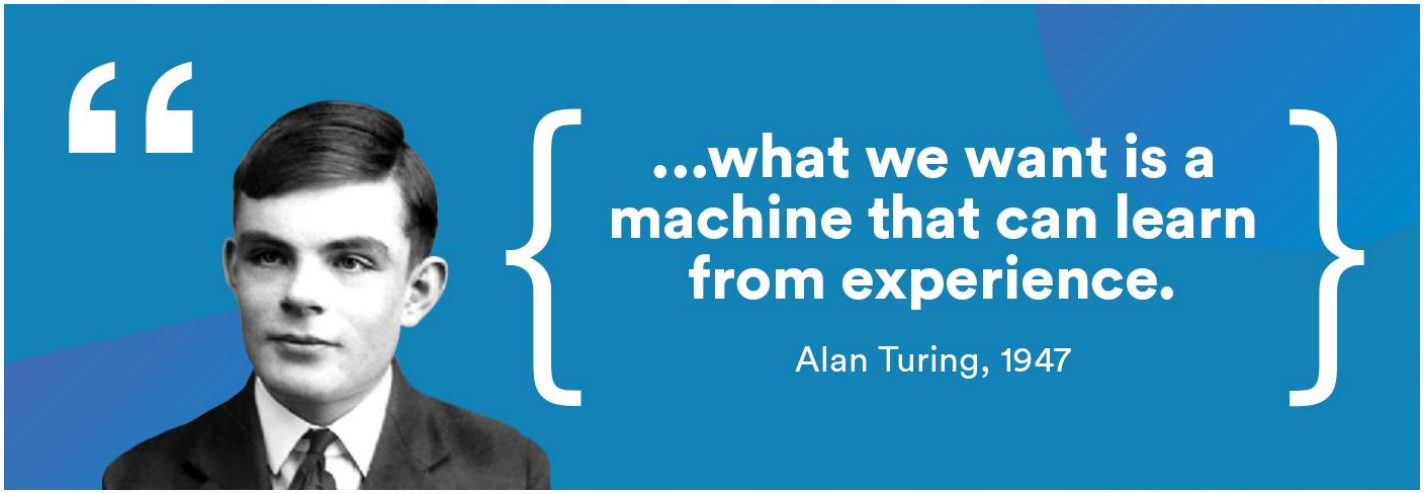
Artifical Intelligence

- Human intelligence exhibited by machines
  - empowers computers to mimic human intelligence such as decision making, text processing, and visual perception.



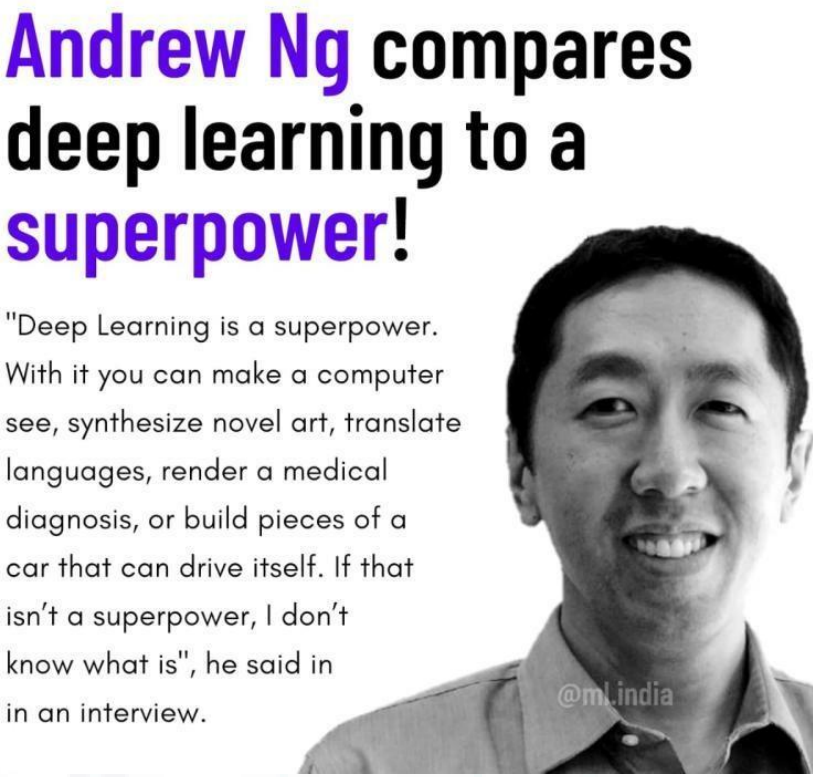
Machine Learning

- An approach to achieve Artificial Intelligence.
  - a subfield of Artificial Intelligence that enables machines to improve at a given task with experience



Deep Learning

- A technique for implementing machine learning
  - a specialized field of Machine Learning that relies on training of Deep Artificial Neural Networks (ANNs) using a large dataset such as images or texts

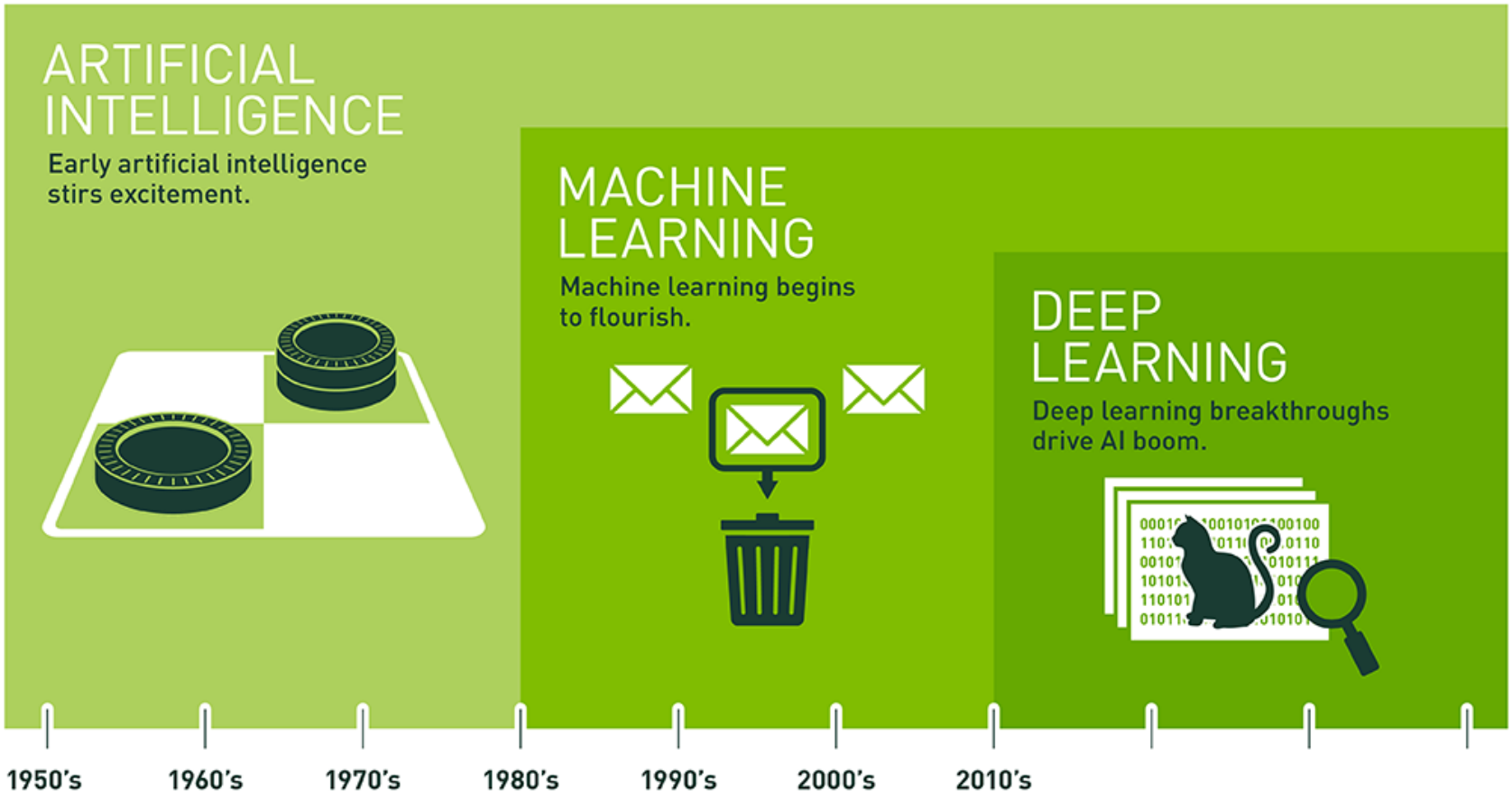


AI vs ML vs DL

- **AI**: Human intelligence exhibited by machines
  - empowers computers to mimic human intelligence such as decision making, text processing, and visual perception.
- **ML**: An approach to achieve Artificial Intelligence.
  - a subfield of Artificial Intelligence that enables machines to improve at a given task with experience
- **DL**: A technique for implementing machine learning
  - a specialized field of Machine Learning that relies on training of Deep Artificial Neural Networks (ANNs) using a large dataset such as images or texts

➔ Thanks to **Deep Learning**, AI has a bright future

AI vs ML vs DL (cont.)



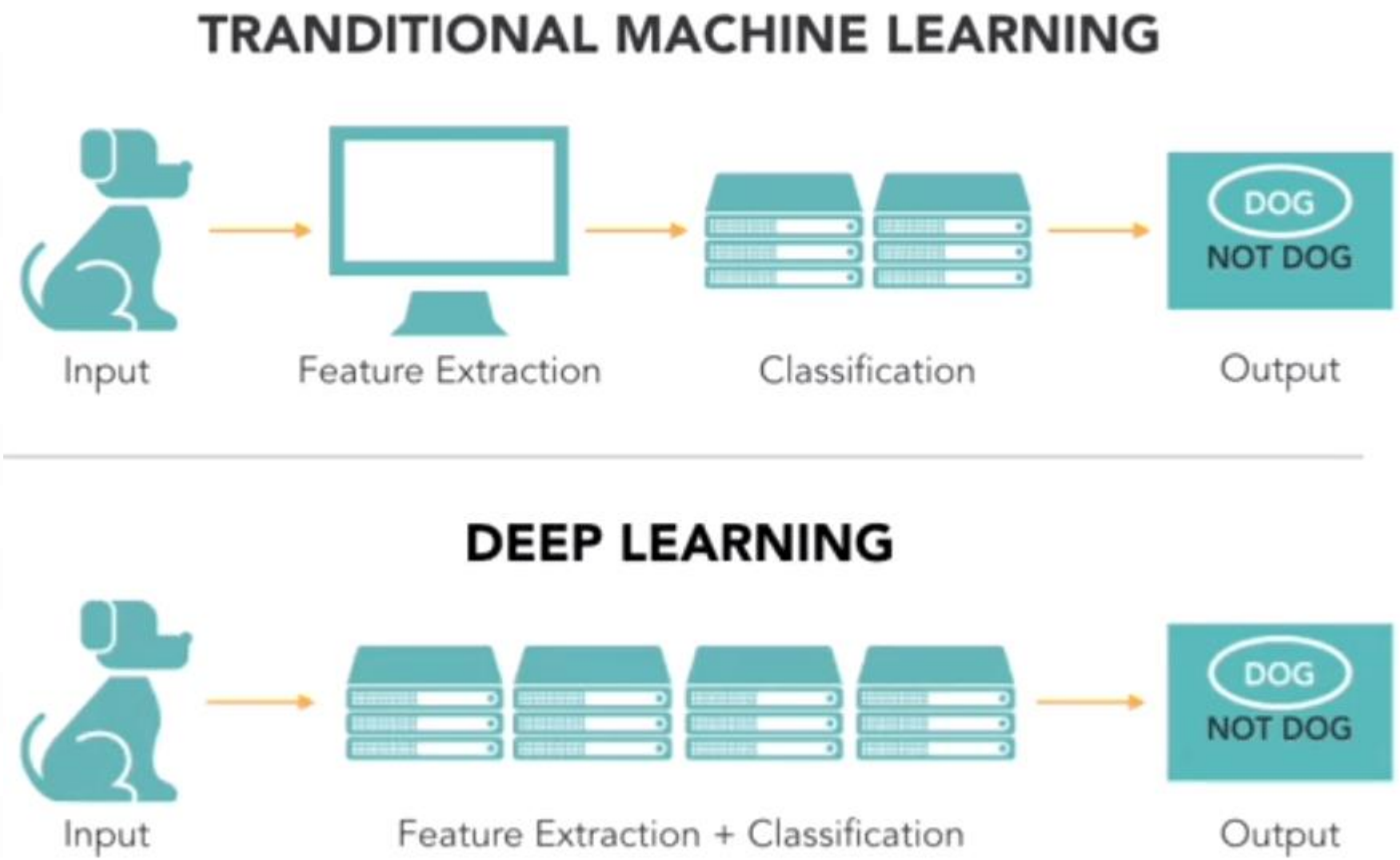
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

ML vs DL

	Machine Learning	Deep Learning
Data Dependencies	Excellent performances on a small/medium dataset	Excellent performance on a big dataset
Hardware dependencies	Work on a low-end machine.	Requires powerful machine, preferably with GPU: DL performs a significant amount of matrix multiplication
Feature engineering	Need to understand the features that represent the data	No need to understand the best feature that represents the data
Execution time	From few minutes to hours	Up to weeks. Neural Network needs to compute a significant number of weights
Interpretability	Some algorithms are easy to interpret (logistic, decision tree), some are almost impossible (SVM, XGBoost)	Difficult to impossible

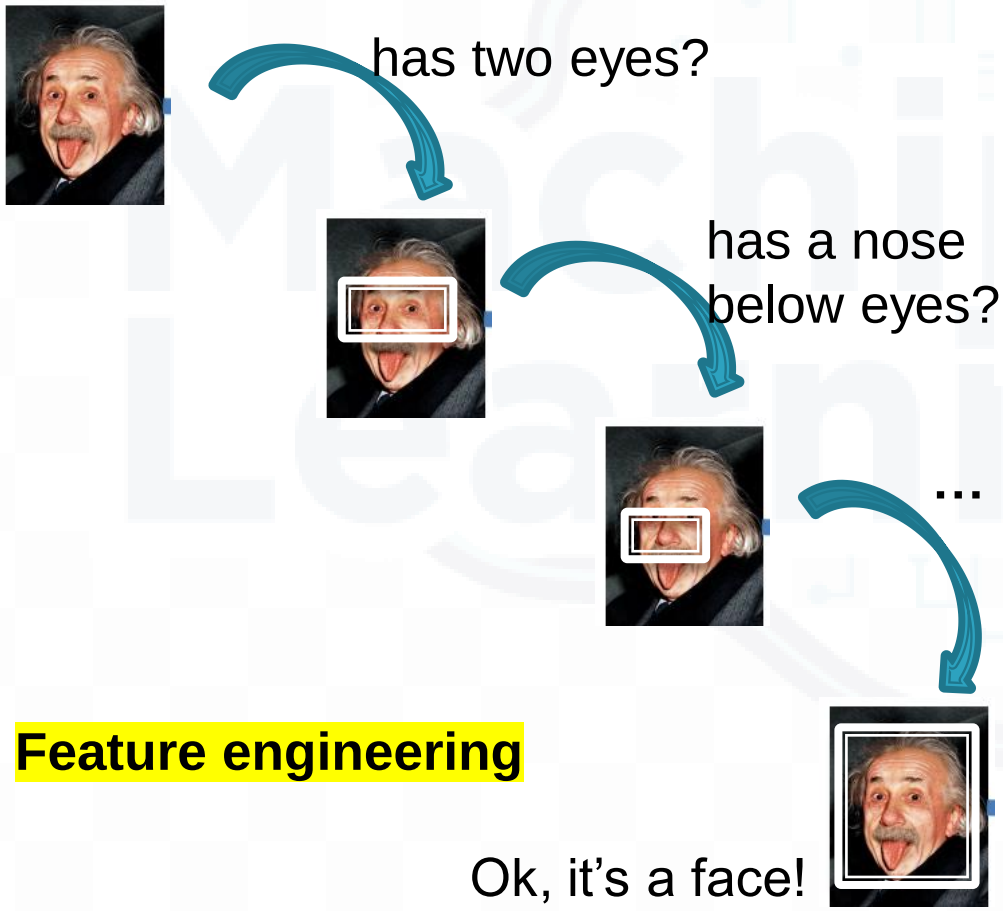
ML vs DL (cont.)

Ex. Classification task



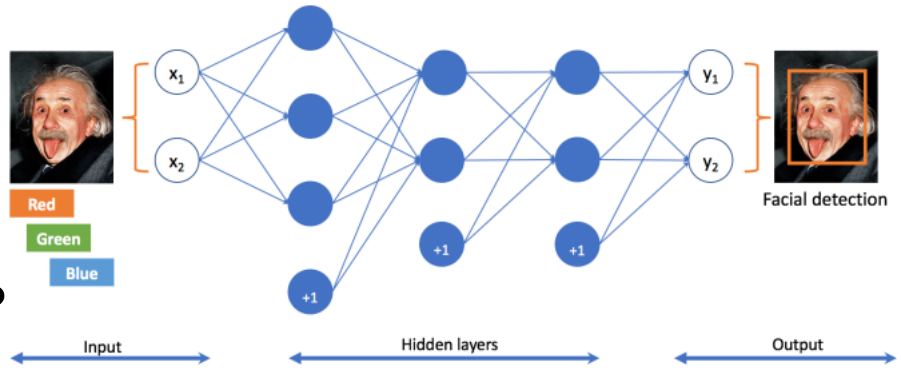
ML vs DL (cont.)

Traditional approach



Feature engineering

Deep learning approach



No feature engineering

ML vs DL (cont.)

- When to use ML or DL?

	Machine learning	Deep learning
Training dataset	Small	Large
Choose features	Yes	No
Number of algorithms	Many	Few
Training time	Short	Long

ML vs Statistics

- Both machine learning and statistics have the same objective
- The same concepts have different names in the two fields

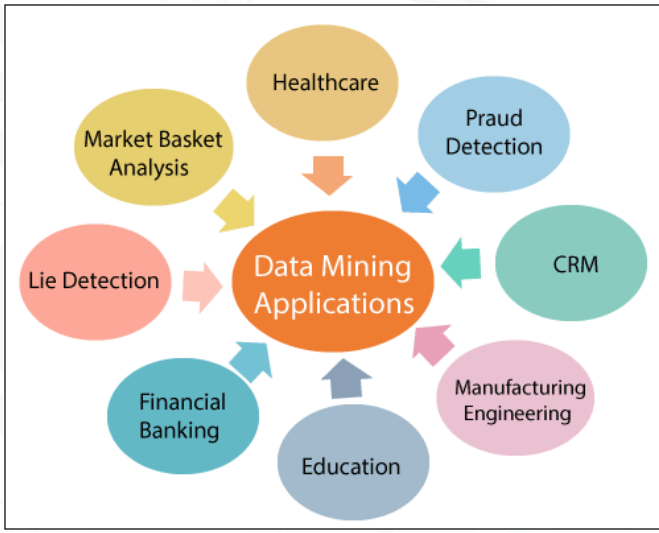
Statistics	Machine Learning
Estimation	Learning
Classifier	Hypothesis
Data Point	Example/ Instance
Regression	Supervised Learning
Classification	Supervised Learning
Covariate	Feature
Response	Label

ML vs Statistic (cont.)

- Machine Learning: a subfield of Computer Science and Artificial Intelligence.
  - Deals with building systems that can learn from data, instead of **explicitly programmed instructions**.
  - A new field.
- Statistic: a subfield of Mathematics.
  - Cheap computing power and availability of large amounts of data allowed data scientists to train computers to learn by analyzing data.
  - Statistical modeling existed long before computers were invented.

Data Mining

- The process of extracting useful information from a huge amount of data.
  - Focused on discovery of previously unknown and important properties in data.
  - Used for extracting patterns from data

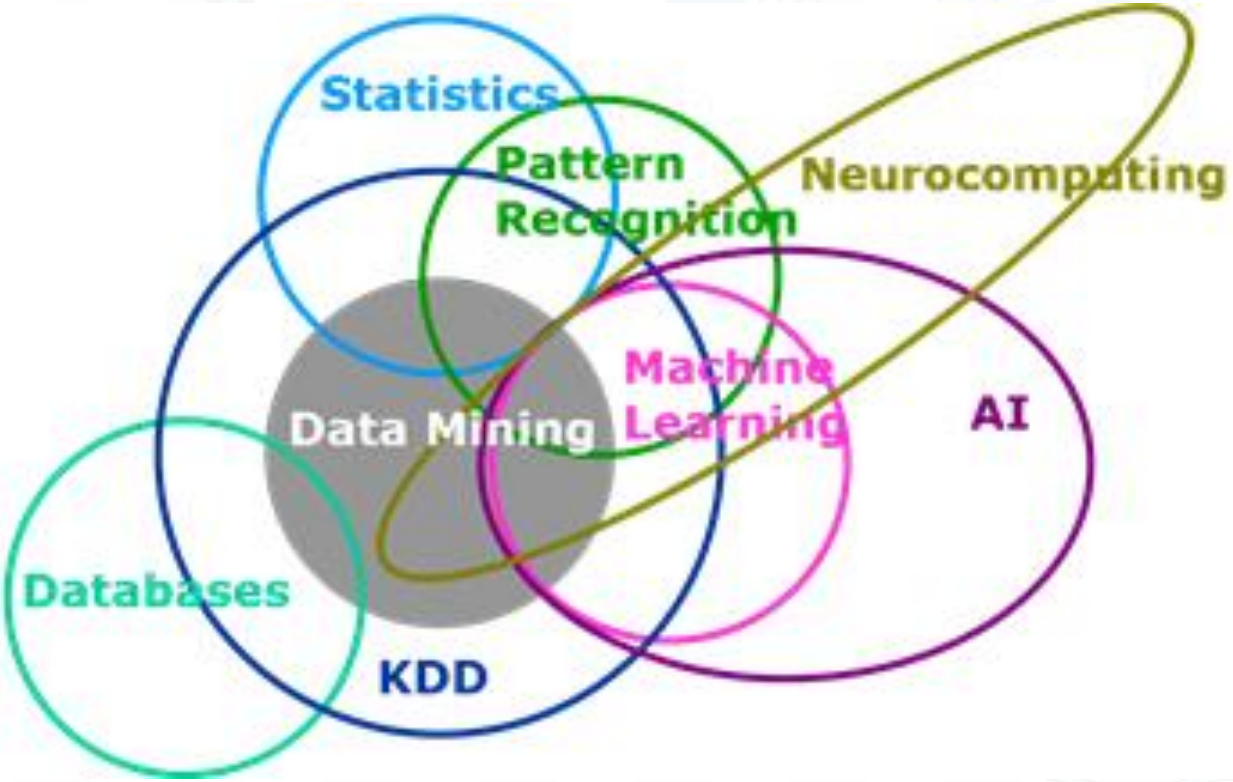


Data Mining (cont.)

- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Control, robotics, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Spam filters, intrusion detection
- Bioinformatics: Motifs, alignment
- Web mining: Search engines
- ...

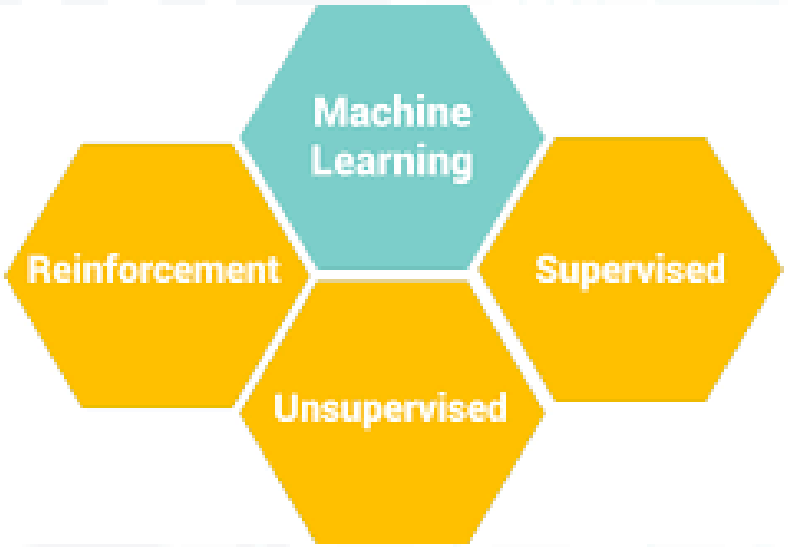
ML vs other Fields

- A Venn diagram that shows how machine learning and statistics are related

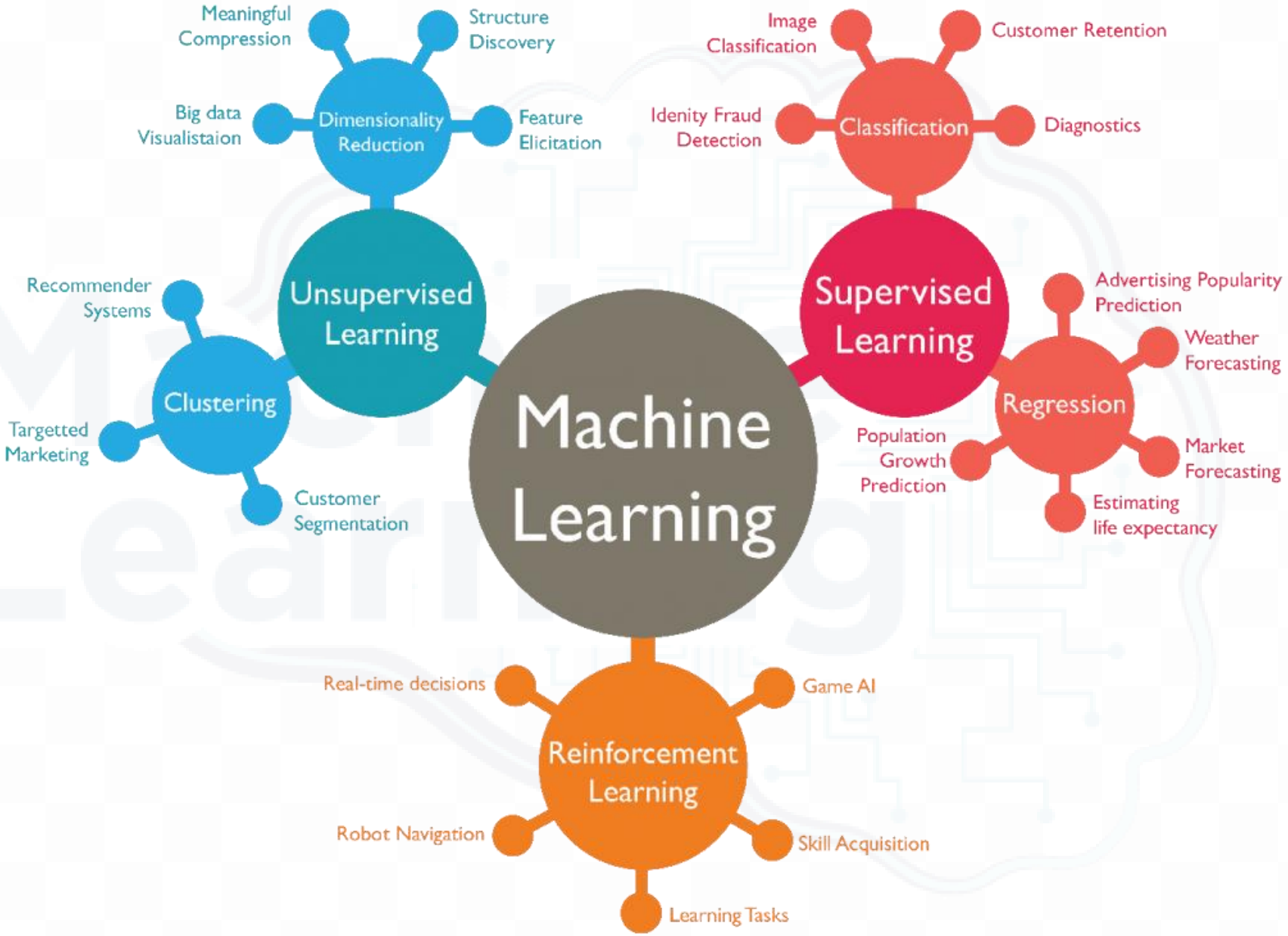


Kdnuggets.com

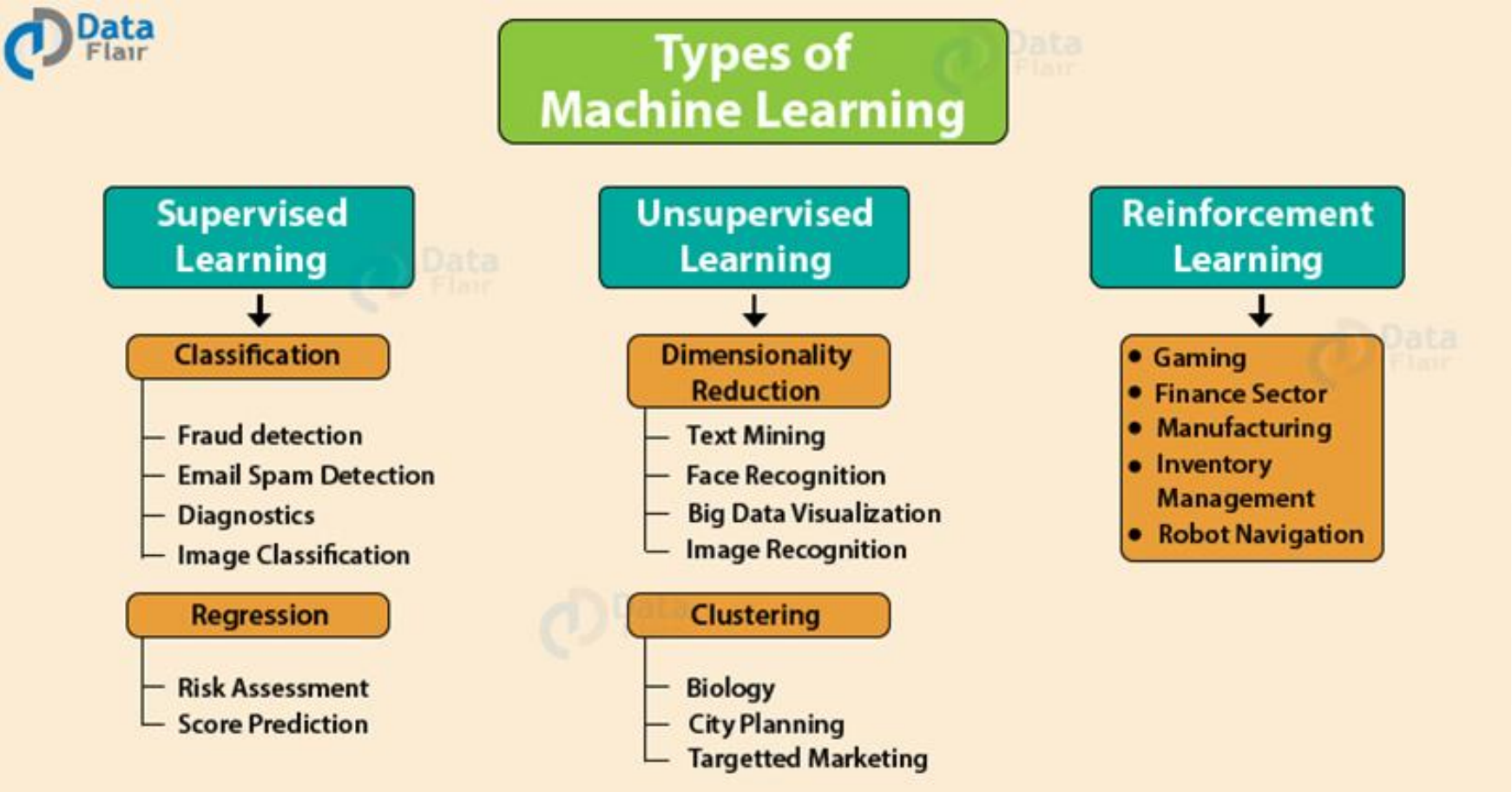
Types of Machine Learning



Types of Machine Learning



Types of Machine Learning (cont.)



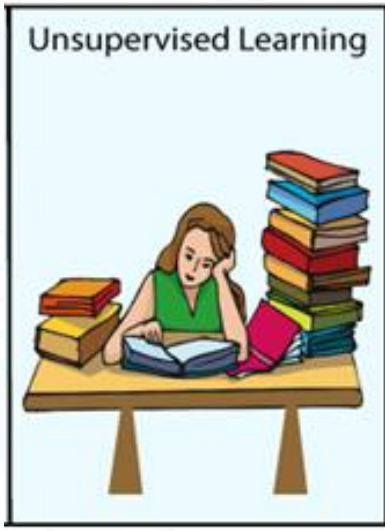
Supervised Learning

- **Supervision**: The training data (i.e., observations, measurements, etc.) are accompanied by labels indicating the class of the observations
  - New data is classified based on the training set

- Example:
  - Spam Detection
    - Map emails to {Spam, Not Spam}
  - Digit recognition
    - Map pixels to {0,1,2,3,4,5,6,7,8,9}
  - Stock Prediction
    - Map new, historic prices, etc. to (the real numbers)

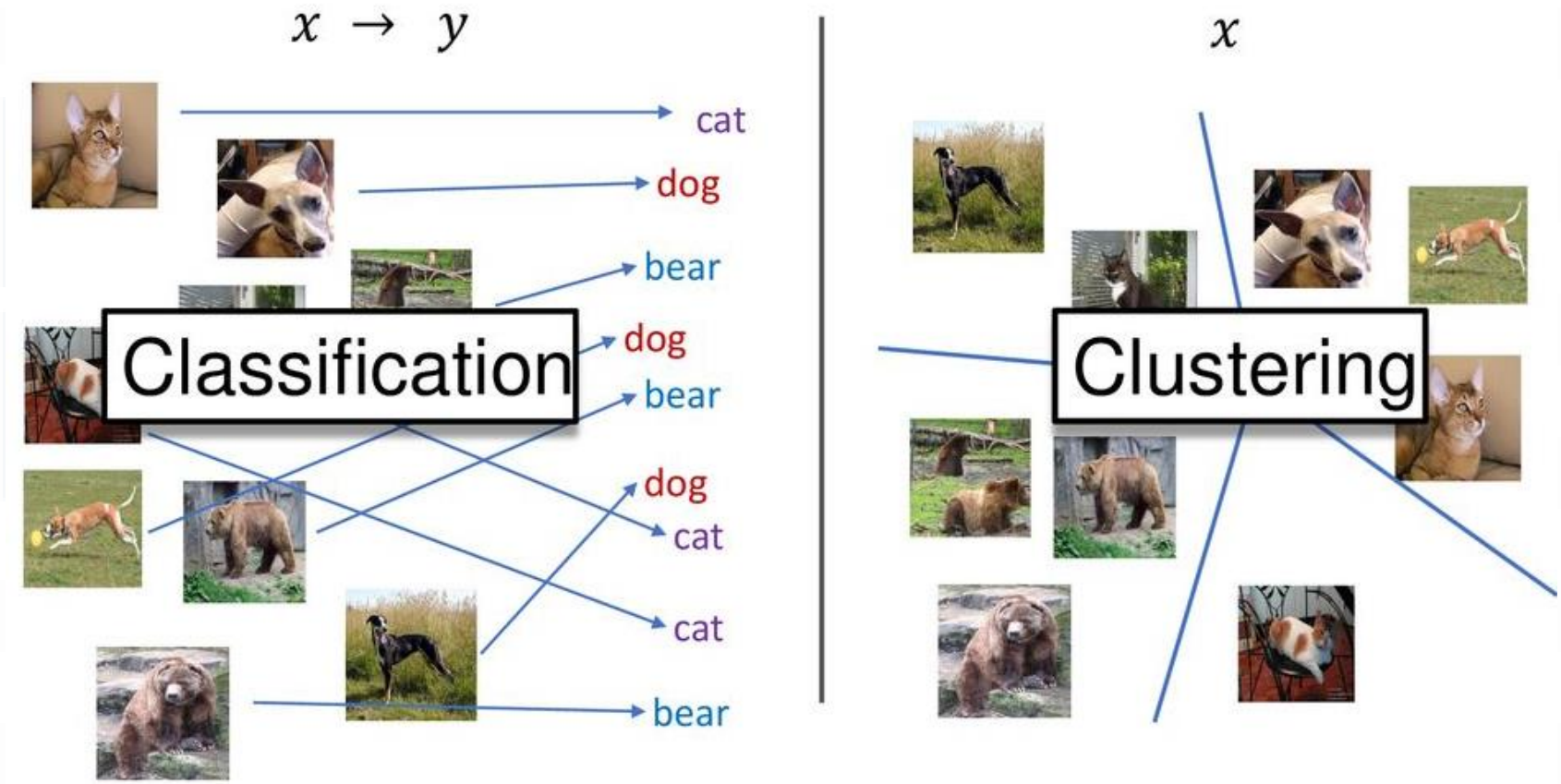
Unsupervised Learning

- The class labels of training data is unknown
- Given a set of measurements, observations, etc. with the aim of establishing the existence of classes or clusters in the data

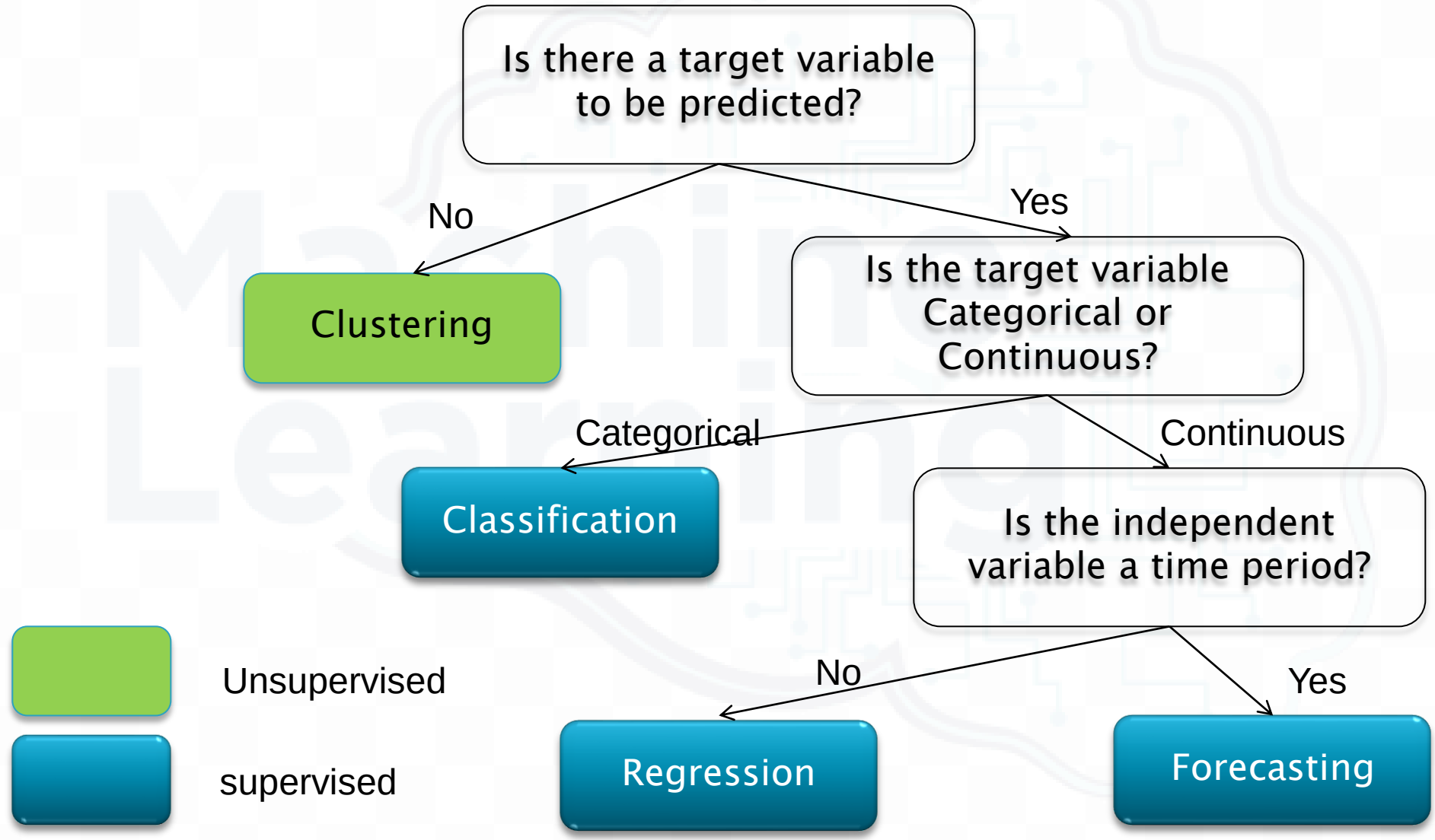


- Example:
  - Customer segmentation in CRM
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs

Supervised vs Unsupervised Learning



Supervised vs Unsupervised Learning



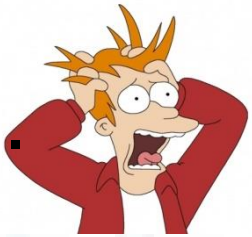
Reinforcement Learning

- Learning a policy: A sequence of outputs

- No supervised output but delayed reward
  - Credit assignment problem
    - the problem of measuring the influence and impact of an action taken by an agent on future rewards
  - Game playing
  - Robot in a maze
  - Multiple agents, partial observability, ...

Reinforcement Learning (cont.)

- Imagine that you were dropped off at an isolated island!
- What would you do?
- Panic? Yes, of course, initially we all would.

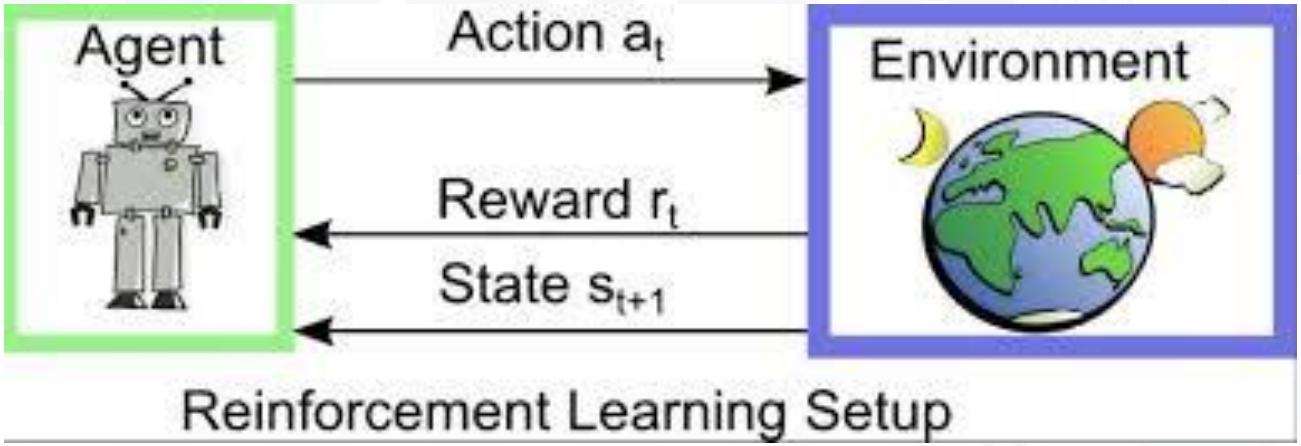


Reinforcement Learning (cont.)

- Later, you will learn how to live on the island by:
  - exploring the environment,
  - understanding the climate condition, the type of food that grows there, the dangers of the island, etc.

Reinforcement Learning (cont.)

- This is exactly how Reinforcement Learning works.
  - an agent is put in an unknown environment
  - Agent must learn by observing and performing actions that result in rewards (maximal).

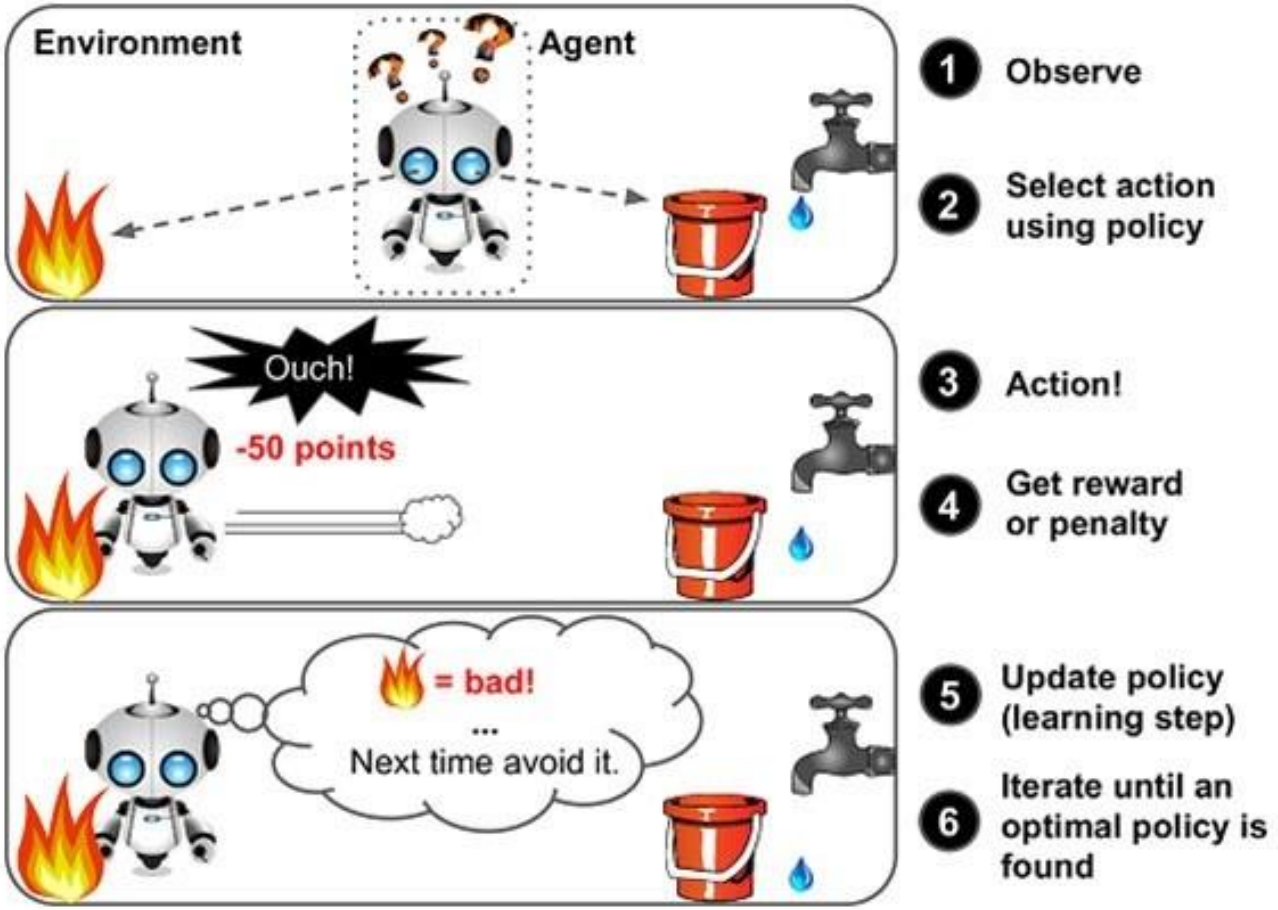


Reinforcement Learning Setup

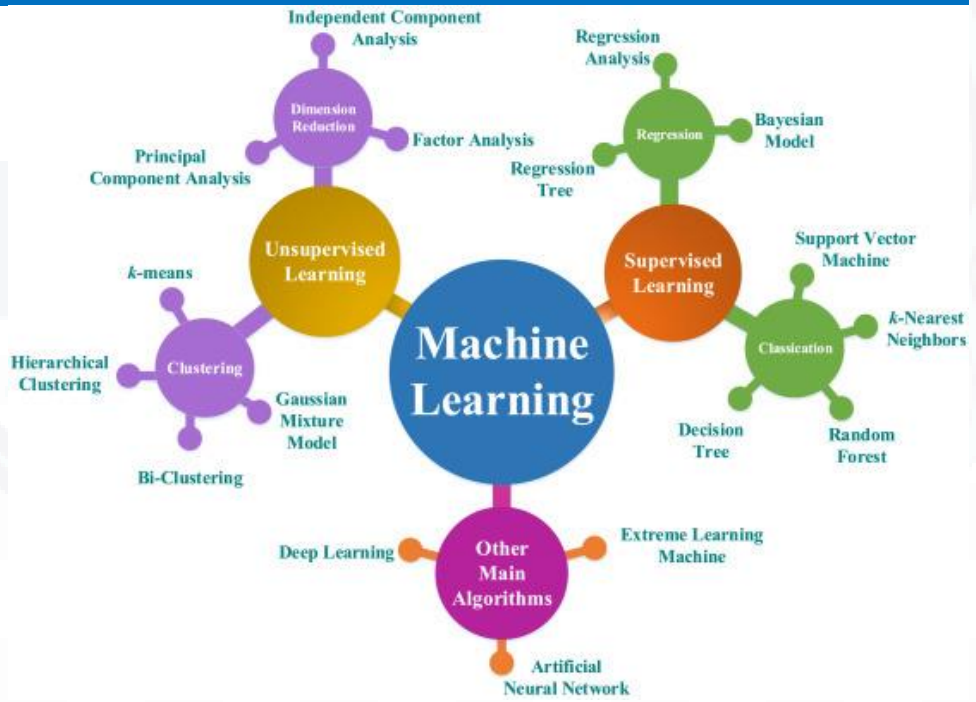
Trial and Error

Reinforcement Learning (cont.)

- An example of Reinforcement Learning:

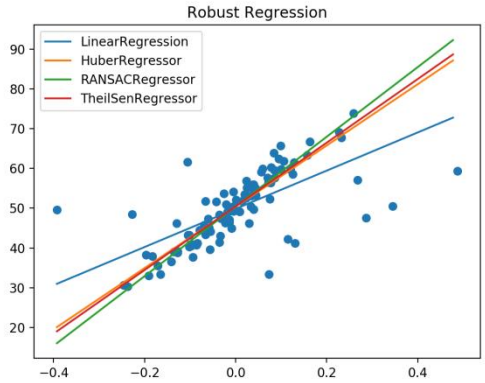


Machine Learning Algorithms

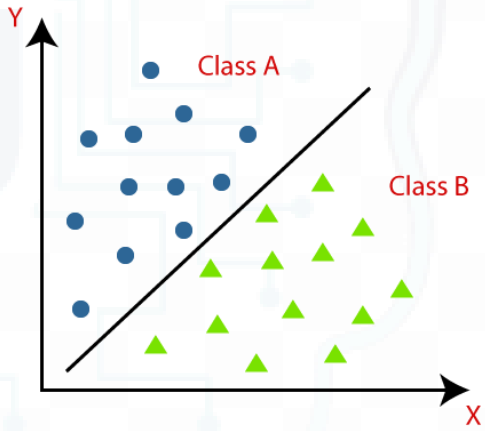


Machine Learning Algorithms

- Regression Algorithms
  - Linear Regression
  - Logistic Regression
  - Stepwise Regression

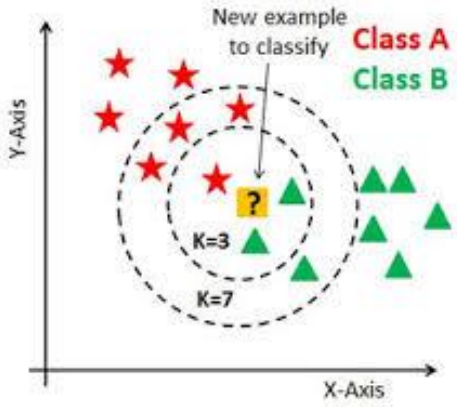


- Classification Algorithms
  - Linear Classifier
  - Support Vector Machine (SVM)
  - Kernel SVM
  - Sparse Representation-based classification (SRC)

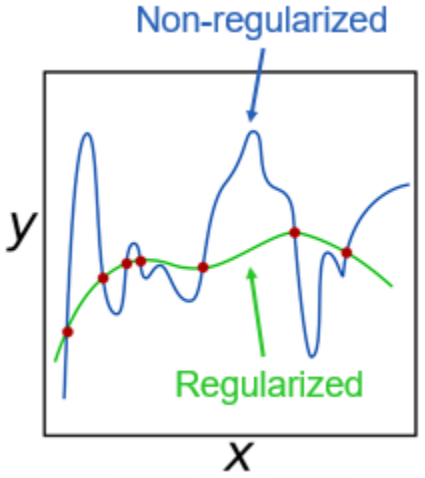


Machine Learning Algorithms (cont.)

- Instance-based Algorithms
  - k-Nearest Neighbor (kNN)
  - Learning Vector Quantization (LVQ)

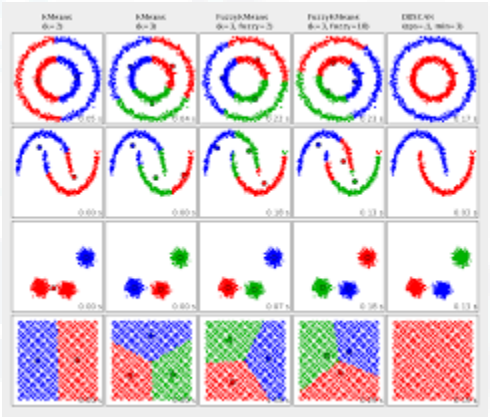


- Regularization Algorithms
  - Ridge Regression
  - Least Absolute Shrinkage and Selection Operator (LASSO)
  - Least-Angle Regression (LARS)
  - Bayesian Algorithms
  - Naive Bayes
  - Gaussian Naive Bayes

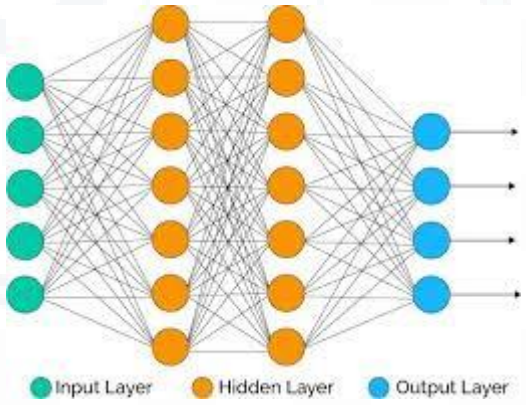


Machine Learning Algorithms (cont.)

- Clustering Algorithms
  - k-Means clustering
  - k-Medians
  - Expectation Maximization (EM)

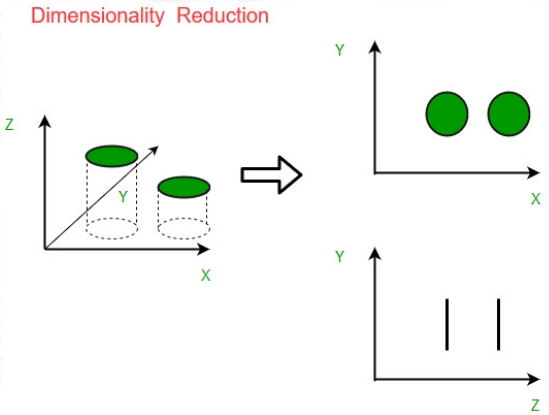


- Artificial Neural Network Algorithms
  - Perceptron
  - Softmax Regression
  - Multilayer Perceptron
  - Backpropagation

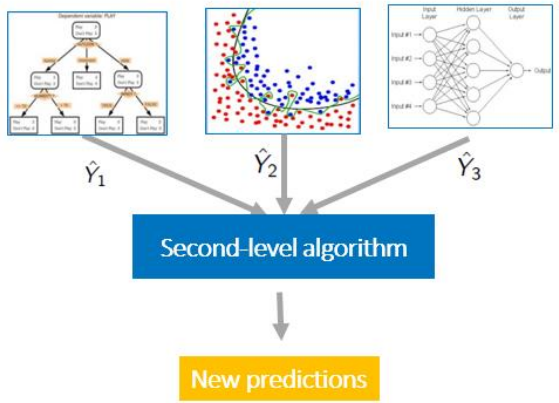


Machine Learning Algorithms (cont.)

- Dimensionality Reduction Algorithms
  - Principal Component Analysis (PCA)
  - Linear Discriminant Analysis (LDA)



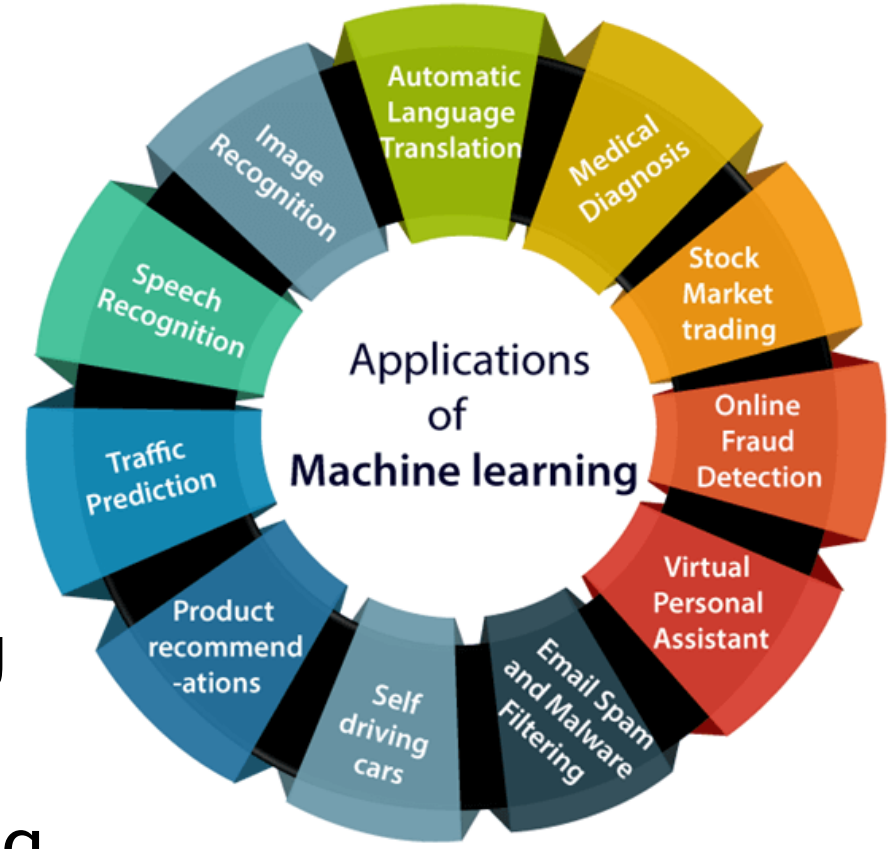
- Ensemble Algorithms
  - Boosting
  - AdaBoost
  - Random Forest



Applications

Applications of ML

- Association
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning



1. Leaning Associations

Learning Associations

- Given a set of records each of which contain some number of items from a given collection
  - $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.
  - Example:
    - $P(\text{chips} \mid \text{beer}) = 0.8$;
    - $P(\text{diaper} \mid \text{beer}) = 0.7$;

Learning Associations (cont.)

- Market-basket analysis
  - Rules are used for sales promotion, shelf management, and inventory management
- Telecommunication alarm diagnosis
  - Rules are used to find combination of alarms that occur together frequently in the same time period
- Medical Informatics
  - Rules are used to find combination of patient symptoms and test results associated with certain diseases

Learning Associations (cont.)

Trang chủ > Máy Ảnh - Máy Quay Phim > Máy ảnh DSLR > Máy Ảnh Canon 750D + Lens 18-55 IS STM (Lê Bảo Minh)





Rê chuột lên hình để phóng to

TIKINOW Máy Ảnh Canon 750D + Lens 18-55 Minh)

★★★★½ (Xem 248 đánh giá)

Đứng thứ 1 trong Top 10 Sacom giảm 1.500.000đ bán chạy tháng này

Thương hiệu: Canon SKU: 6001759234016

DEAL HOT Giá: 11.890.000 ₫

0% TRẢ GÓP Chỉ 990.833 ₫/tháng trong 12 tháng qua thẻ tín dụng (?)

Tiết kiệm: 32% (5.610.000 ₫)

Giá thị trường: 17.500.000 ₫

Khuyến mãi kết thúc sau 2 ngày 11 : 30 : 47

Đã bán 389

- Cảm biến: CMOS 24.2MP
- Bộ xử lý hình ảnh: DIGIC 6

SẢN PHẨM THƯỜNG ĐƯỢC XEM CÙNG

tiki.vn

Canon M10 KIT 15-45mm (Lê Bảo Minh)

5.900.000 ₫ ~~10.500.000 ₫~~ -44%

Máy Ảnh Sony Alpha A6000 + 16-50mm

11.790.000 ₫ ~~14.990.000 ₫~~ -21%

Máy Ảnh Sony Alpha A6300 + Kit 16-50mm - Đen

19.490.000 ₫ ~~23.990.000 ₫~~ -19%

Lens Canon 50mm f/1.8 STM (Lê Bảo Minh)

2.500.000 ₫ ~~2.990.000 ₫~~ -16%

Máy Ảnh Sony Alpha A6000 (Body)

9.690.000 ₫ ~~12.490.000 ₫~~ -22%

Learning Associations (cont.)

Click để xem bìa sau

Cảm Ơn Người Lớn (Bìa Cứng) (Kèm Quà Tặng - Số Lượng Có Hạn)

Tác giả: Nguyễn Nhật Ánh

Nhà xuất bản: Nxb Trẻ

Nhà phát hành: NXB Trẻ

Thích Chia sẻ 2 người thích nội dung này. Hãy là người đầu tiên trong số bạn bè của bạn.

★★★★☆ (1 đánh giá . 1 bình luận)

Thông tin kèm theo

Miễn phí giao hàng toàn quốc cho Đơn hàng từ 250.000đ (Áp dụng từ 1/2/2015. Xem chi tiết »)

Phiên bản khác đang bán

Cảm Ơn Người Lớn (Bìa Mềm) (Kèm Quà Tặng ...
Nguyễn Nhật Ánh
★★★★★ 110.000 đ
88.000 đ

Sách nên mua kèm với sách này

Tổng giá bán 122.000 ₫

MUA NGAY

Thông tin thanh toán

Giá bìa	~~220.000 ₫~~
Giá bán	154.000 ₫
Tiết kiệm	66.000 ₫ (30%)
Chất lượng sách	Loại A (?)

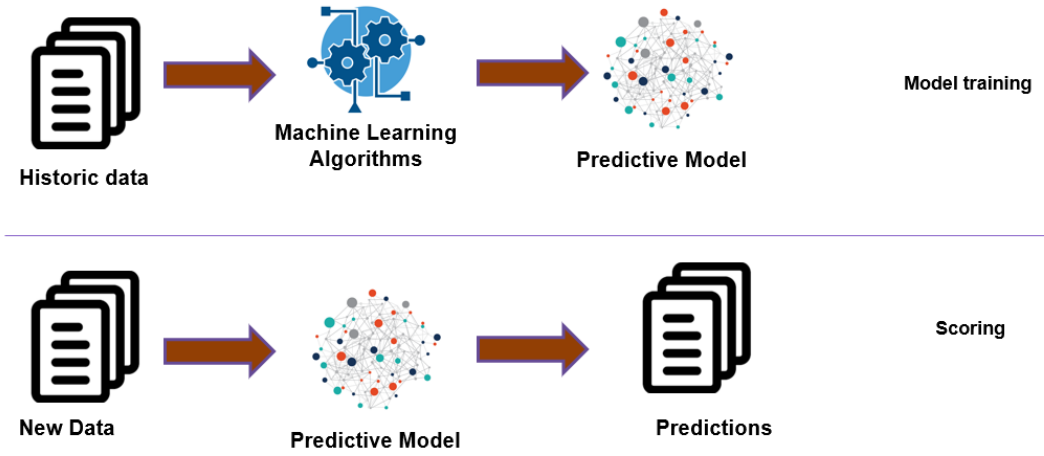
SẮP HẾT HÀNG

MUA NGAY

THÍCH ĐÃ ĐỌC

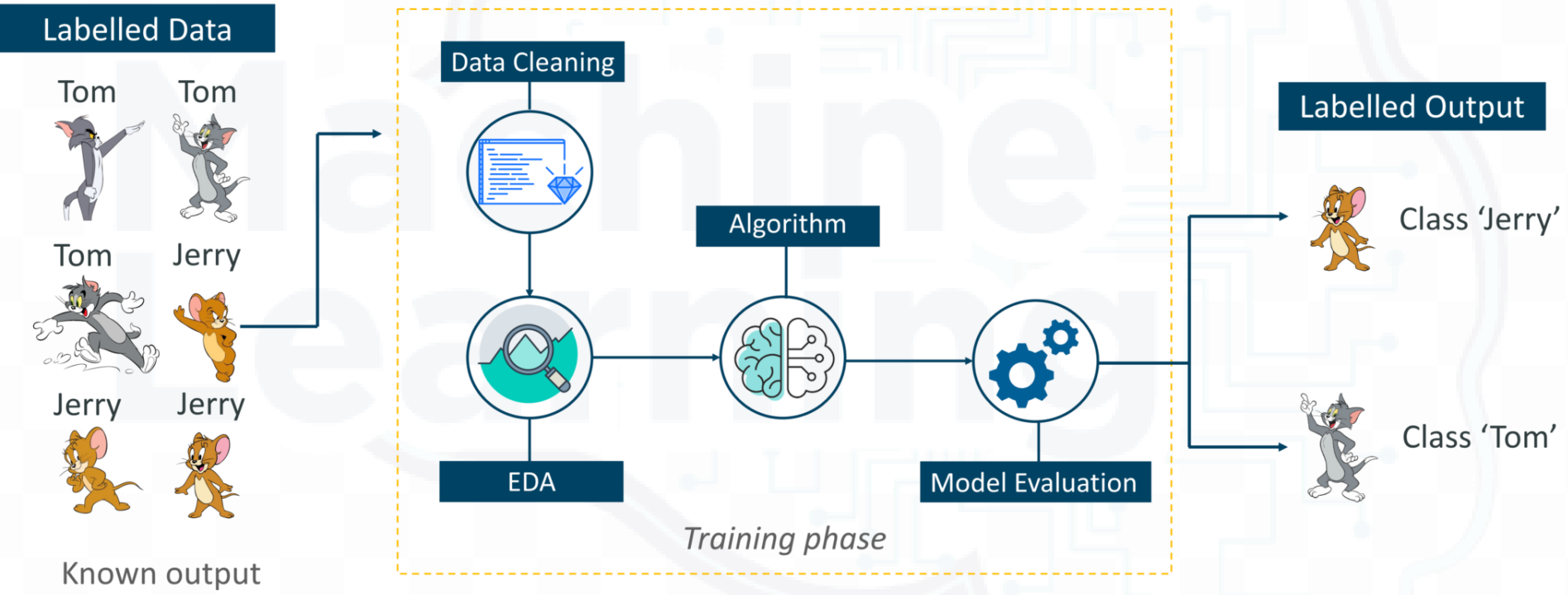
vinabook.com

2. Supervised Learning



Classification

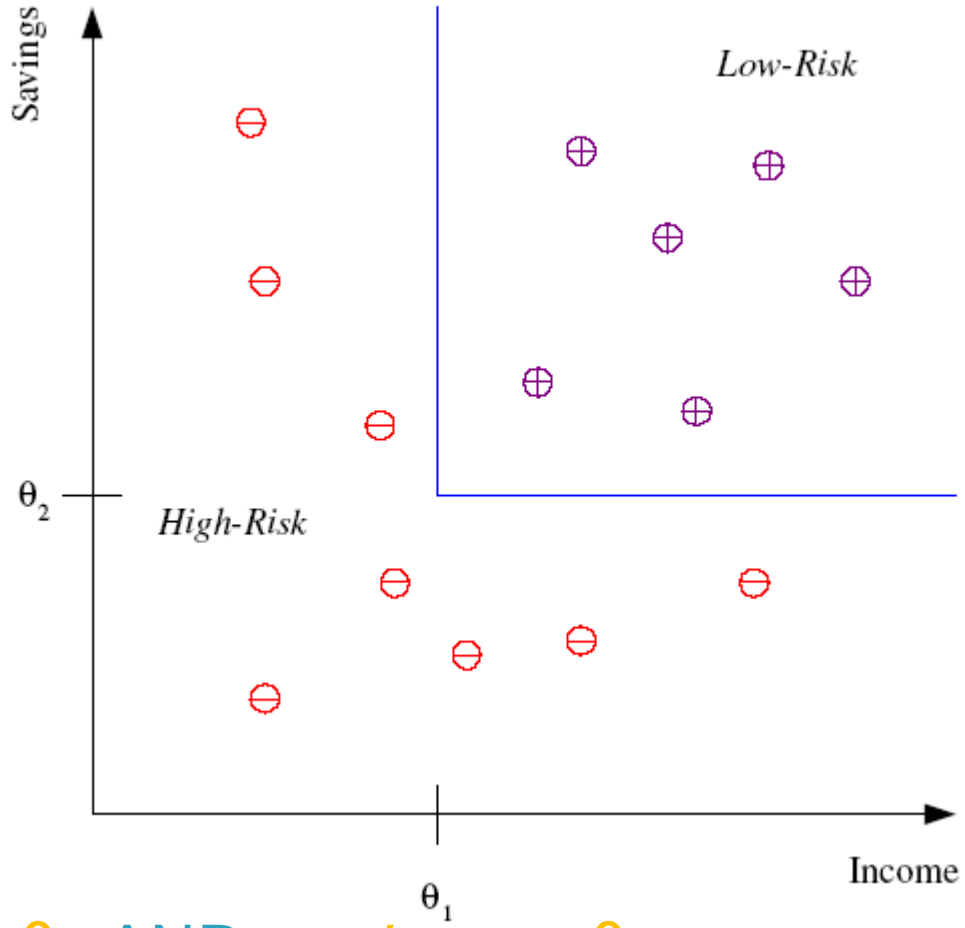
- How does classification work?



EDA: Exploratory Data Analysis

Classification (cont.)

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* > θ_1 AND *savings* > θ_2
THEN low-risk ELSE high-risk

Classification (cont.)

- Aka Pattern recognition
- Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style
- Character recognition: Different handwriting styles.
- Speech recognition: Temporal dependency.
- Medical diagnosis: From symptoms to illnesses
- Biometrics: Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc
- ...

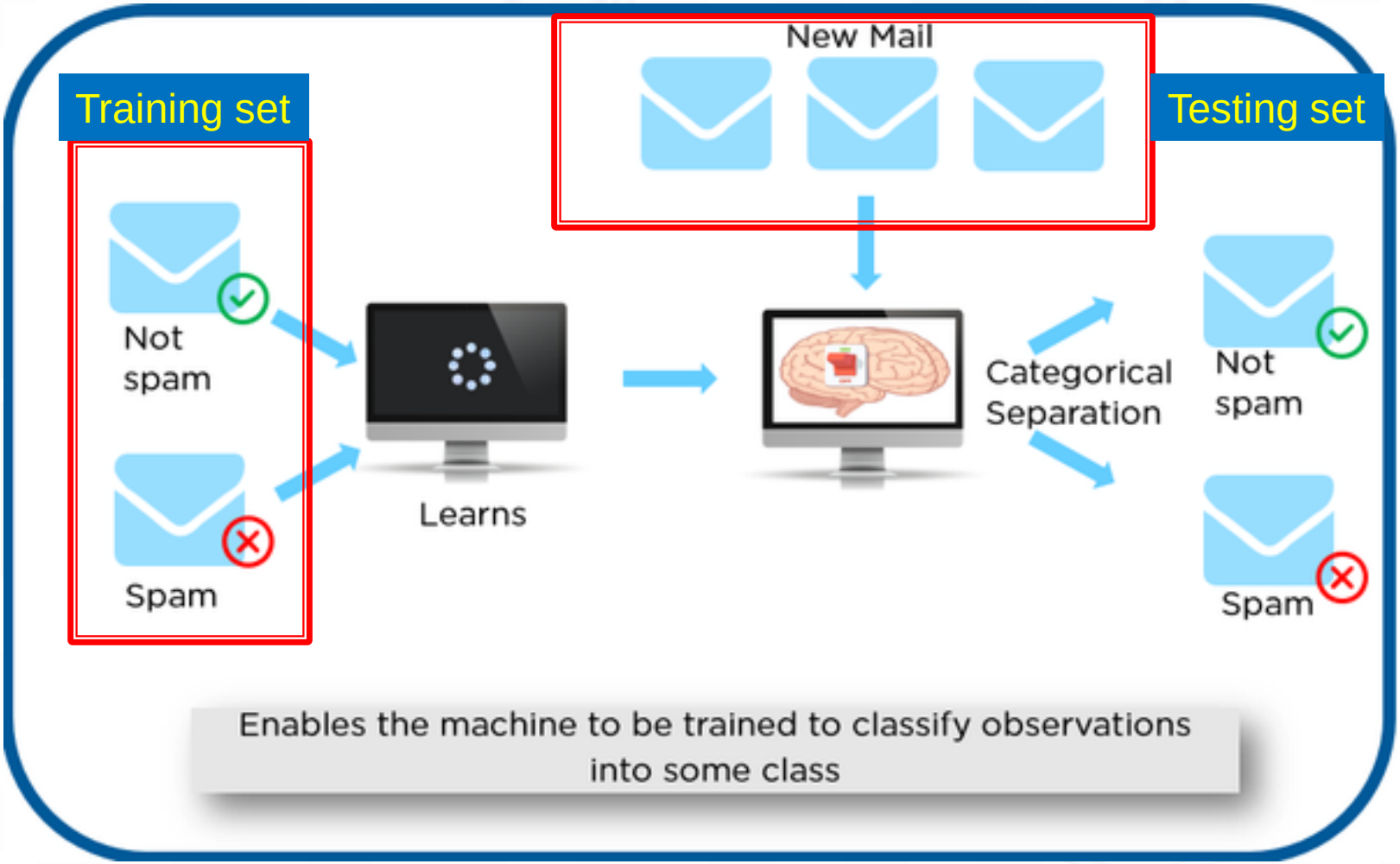
Face Recognition

Training examples of a person

Test images

ORL dataset,
AT&T Laboratories, Cambridge UK

Spam mail detection

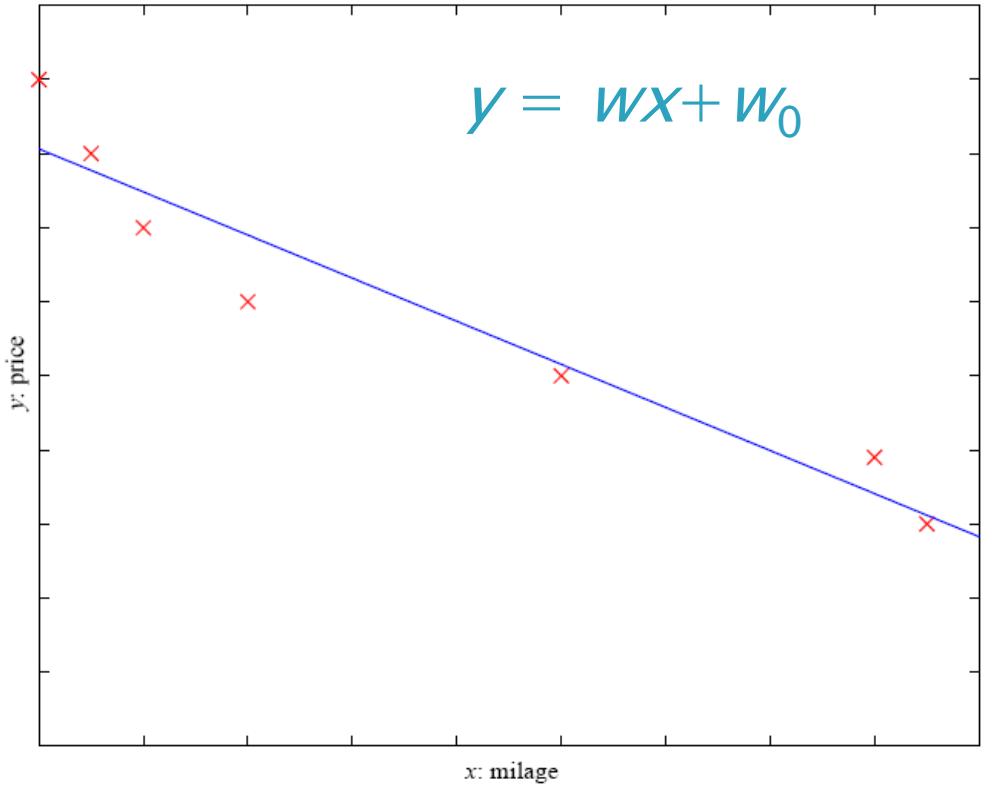


Regression

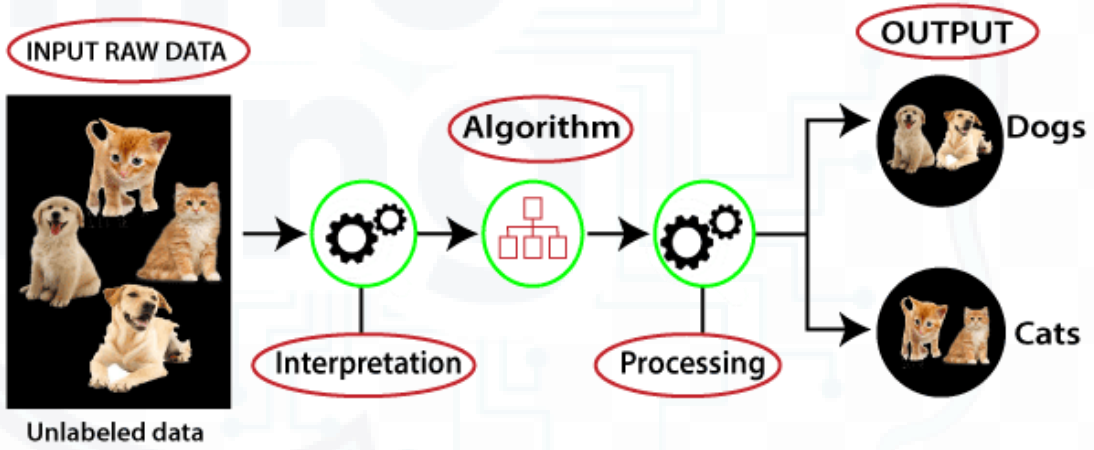
- Example: Price of a used car
  - x : car attributes
  - y : price

$y = g(x \mid \theta)$

$g(\,)$ model,
θ parameters

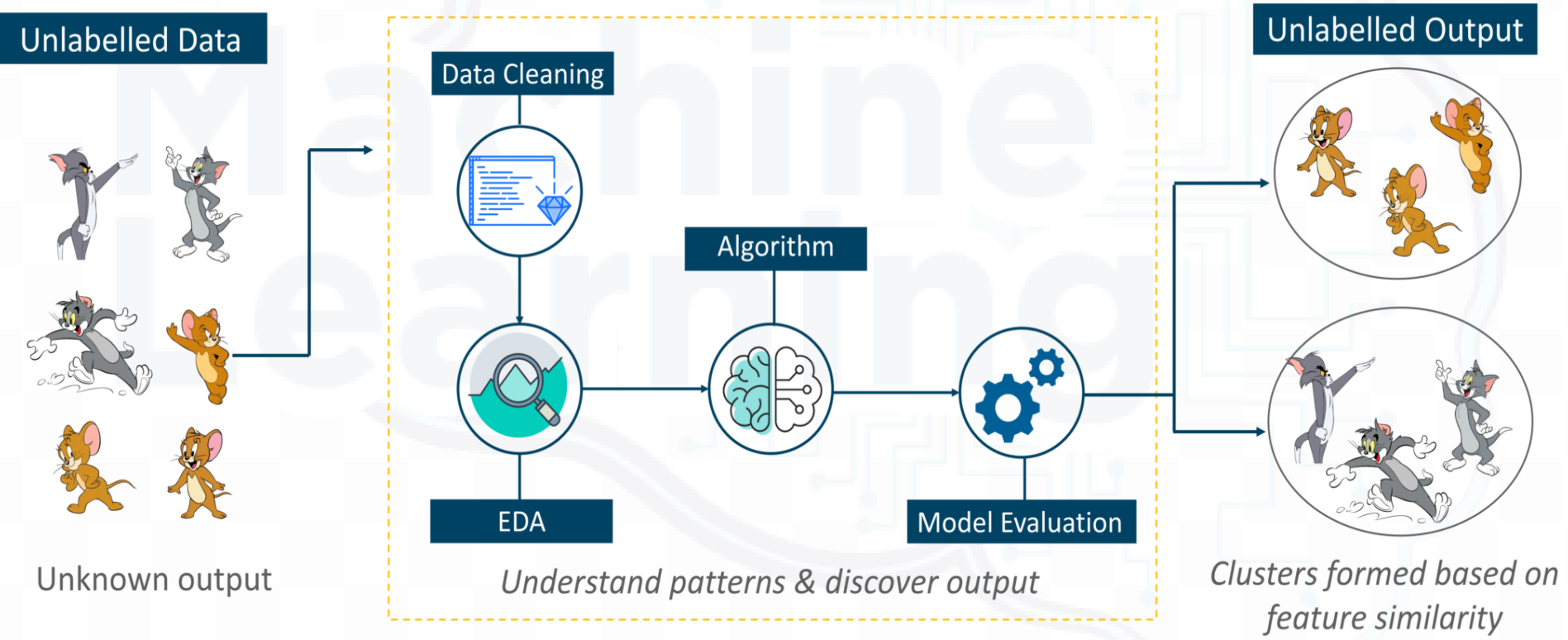


3. Unsupervised Learning



Clustering

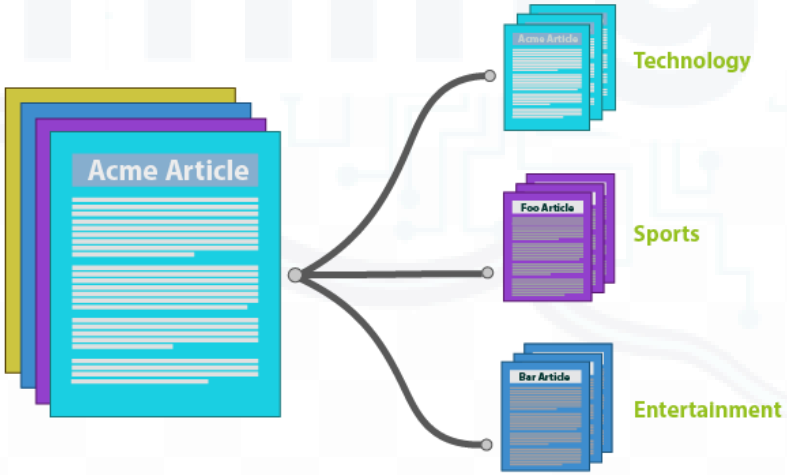
- How does clustering work?



EDA: Exploratory Data Analysis

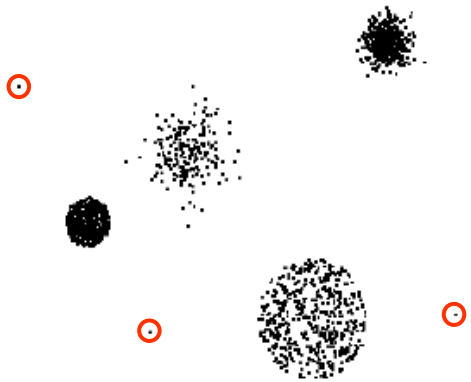
Clustering: Application

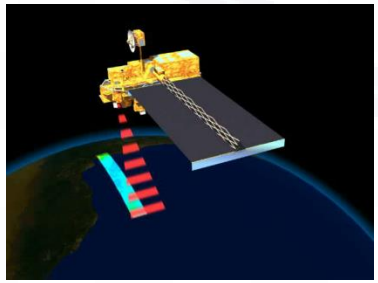
- Document Clustering:
  - **Goal**: To find groups of documents that are similar to each other based on the important terms appearing in them.
  - **Approach**: To identify frequently occurring terms in each document. Form a similarity measure based on the frequencies of different terms. Use it to cluster.

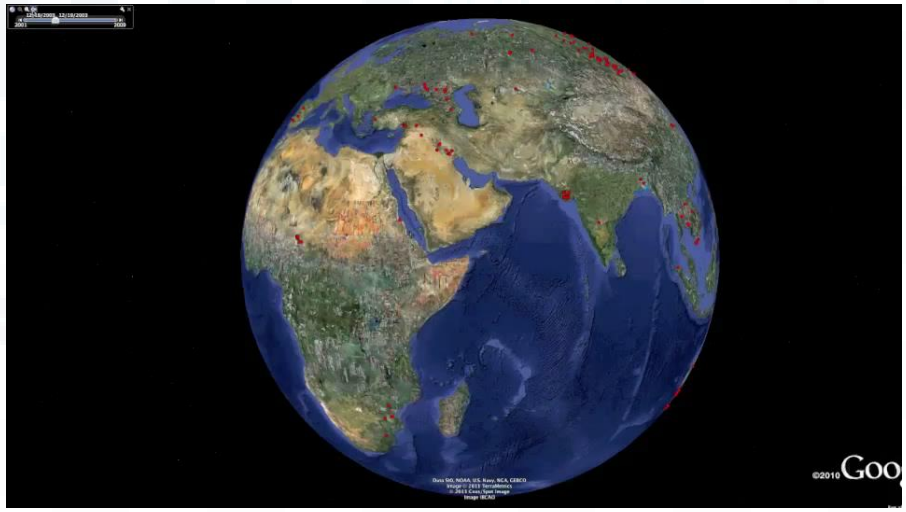


Deviation/Anomaly/Change Detection

- Detect significant deviations from normal behavior
- Applications:
  - Credit Card Fraud Detection
  - Network Intrusion Detection
  - Identify anomalous behavior from sensor networks for monitoring and surveillance.
  - Detecting changes in the global forest cover.

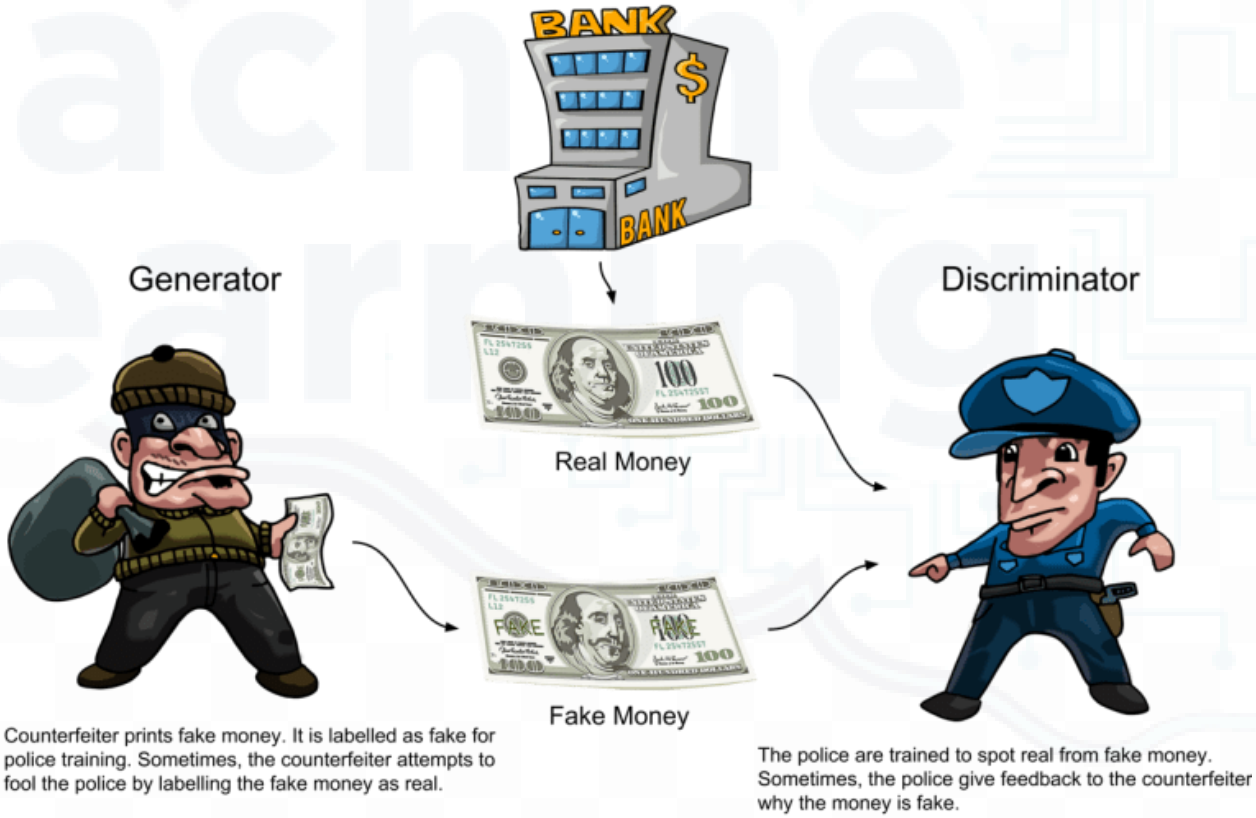






GAN: Generative Adversarial Networks

- GAN: learn to model the input distribution by training two competing (and cooperating) networks called **generator** and **discriminator**

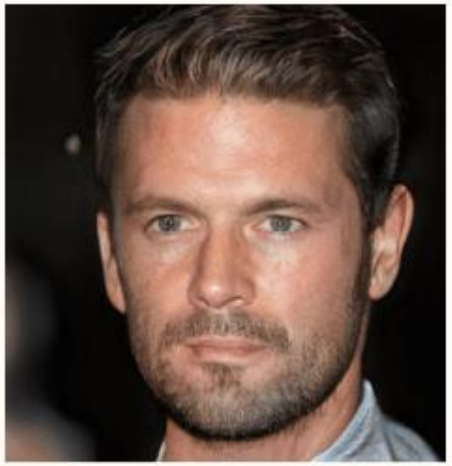


GAN (cont.)

2014

2015

2016



2017

2018

GAN (cont.)

- Generate Anime characters

Towards the Automatic Anime Characters Creation with Generative Adversarial Networks

GAN (cont.)

- Text to image

StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks

GAN (cont.)

- Music generation

(a) MidiNet model 1

(b) MidiNet model 2

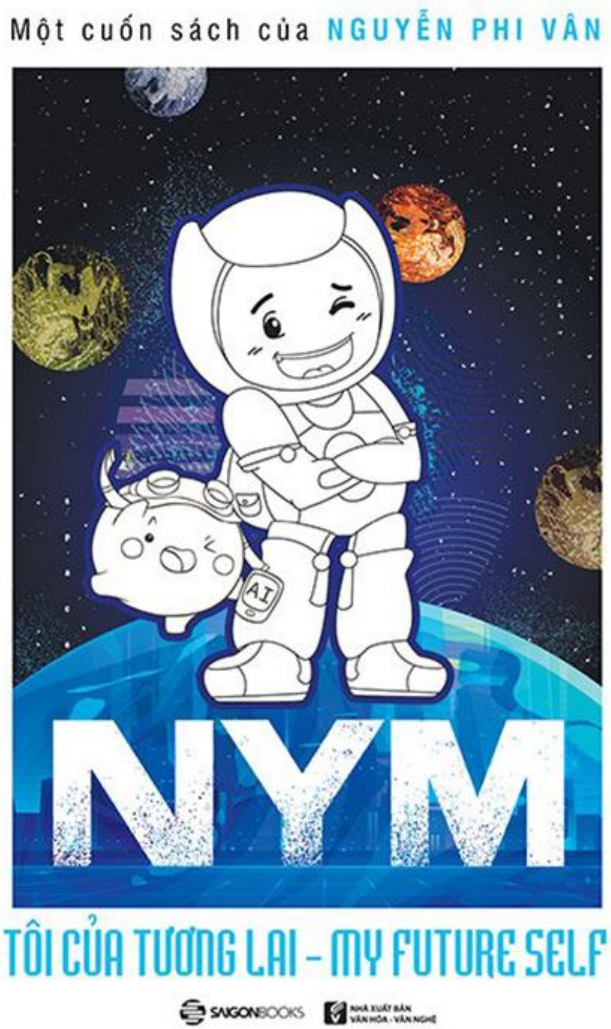
(c) MidiNet model 3

Figure 3. Example result of the melodies (of 8 bars) generated by different implementations of MidiNet.

MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation

AI for Writing

- **Nym**: AI do Nguyễn Phi Vân xây dựng trong hơn ba năm, được nạp dữ liệu về ngôn ngữ, kiến thức đời sống.
  - AI đã nói chuyện trực tiếp, một – một với 11 triệu người trẻ Việt trên facebook để thu thập thông tin.
  - **Nym – Tôi của tương lai** – quyển sách con người và AI hợp tác viết.



AI news anchor

- 3 AI news anchors appeared on Indonesian television

AI singer

- Ann:
  - developed by Bobo Đặng, 3/2023
  - MV "Làm sao nói thương anh"

ChatGPT

- ChatGPT:
  - a deep learning language model
  - developed by OpenAI,
  - which is capable of generating human-like text based on the input provided

OPENAI: Sora

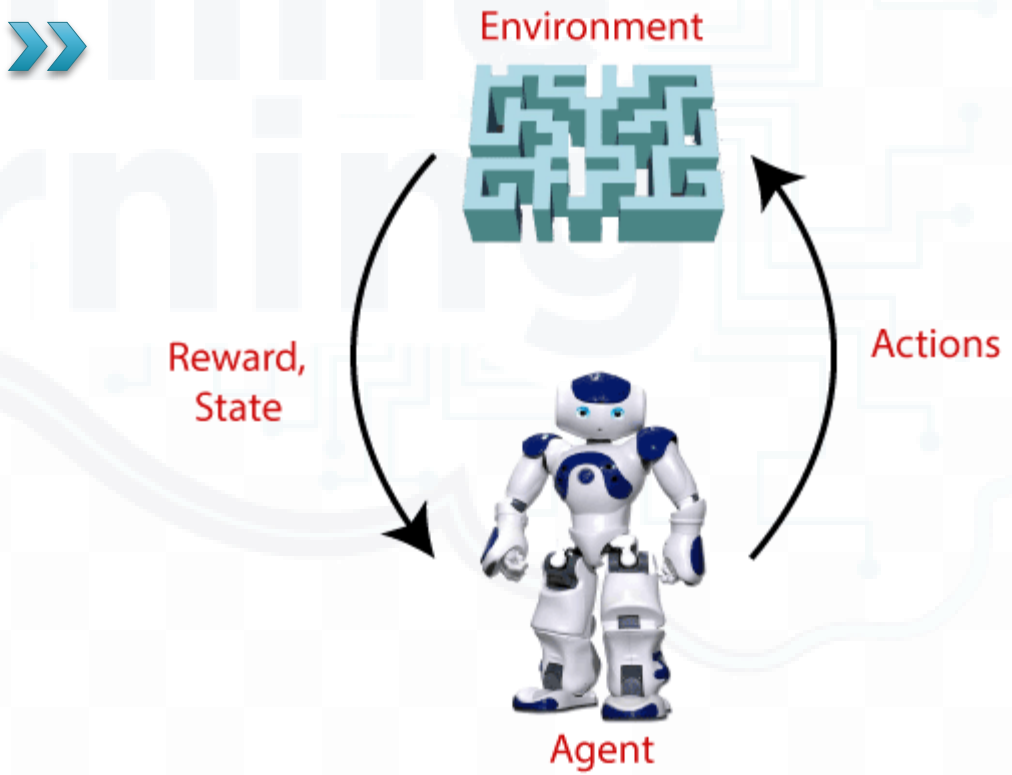
Sora (openai.com)

- Sora is an AI model that can create realistic and imaginative scenes from text instructions.
  - text-to-video model
  - can generate videos up to a minute long while maintaining visual quality and adherence to the user's prompt.

OPENAI: Sora

- **Prompt**: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

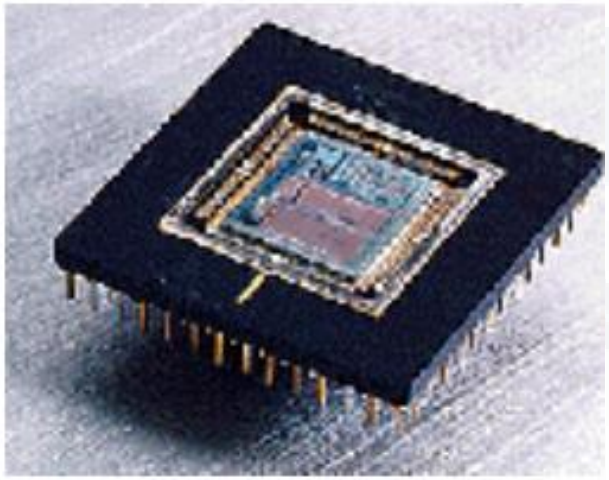
4. Reinforcement Learning



Game Playing – Chess

- 1997, Deep Blue (IBM) has defeated Garry Kasparov

won 3 games,
lost 2,
tied 1



Deep Blue chip

Game Playing – Go

- 2016, AlphaGo (Deepmind, Google) has defeated Lee Sedol.

won 4 games,
lost 1

AlphaZero

Robotic Vehicles

- 2006, a driverless robotic car – STANLEY won in DARPA Grand Challenge (dessert)

- 2007, a driverless robotic car – CMU's BOSS won in Urban Challenge

FACULTY OF INFORMATION TECHNOLOGY