Machine Learning
Lecture 1: Introduction

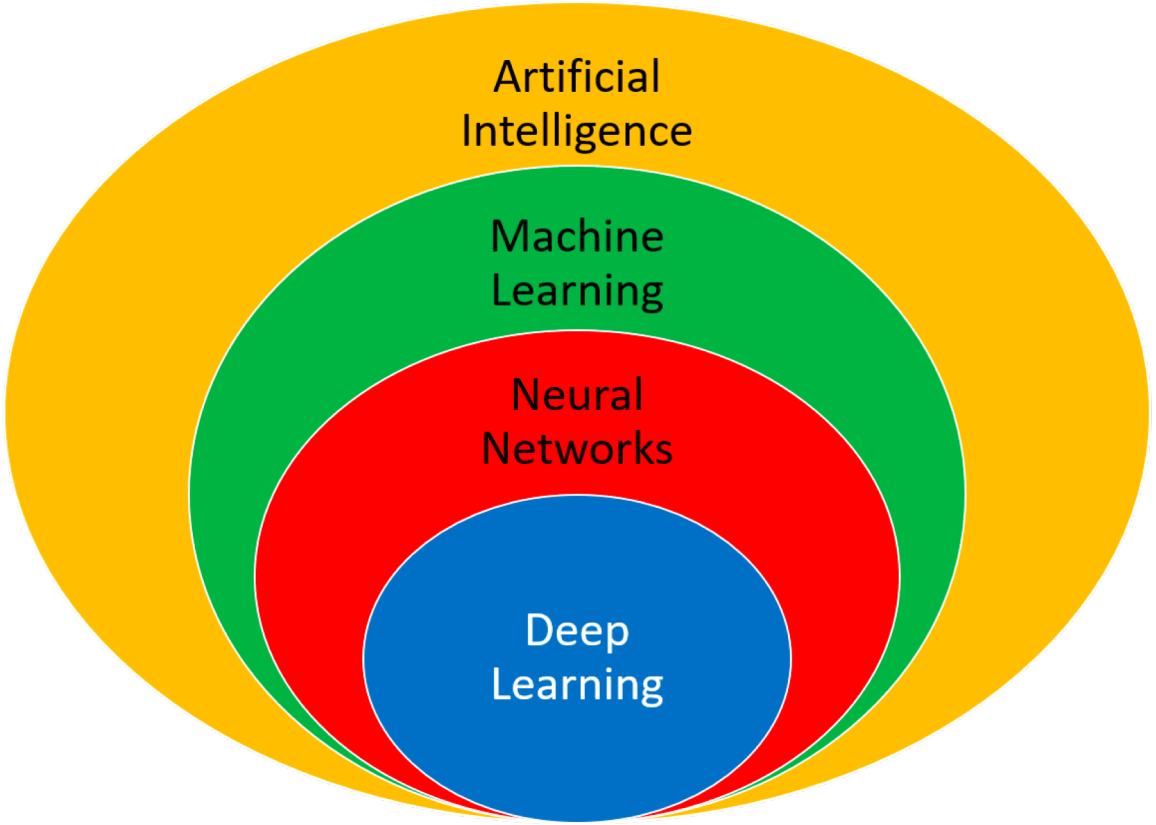
LÊ ANH CƯỜNG

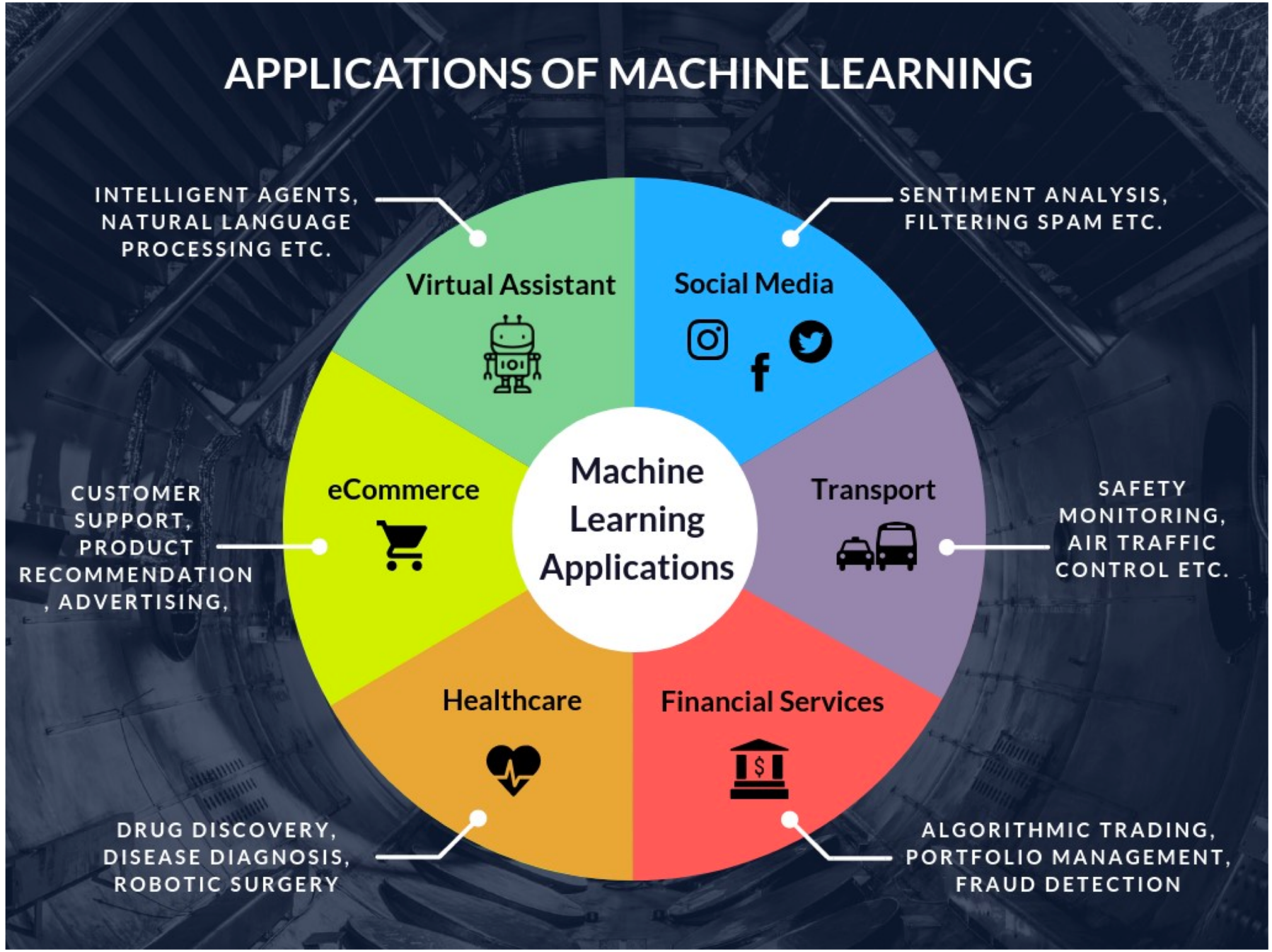
Tôn Đức Thắng University (TDTU)

Outline

1. Machine learning and Artificial Intelligence
2. A general paradigm for AI systems and the ML component
3. Examples of ML models
4. Categorization of Machine learning problems and methods

AI, Machine Learning, and Deep Learning




APPLICATIONS OF MACHINE LEARNING
INTELLIGENT AGENTS, NATURAL LANGUAGE PROCESSING ETC.
Virtual Assistant
SENTIMENT ANALYSIS, FILTERING SPAM ETC.
Social Media
CUSTOMER SUPPORT, PRODUCT RECOMMENDATION , ADVERTISING,
eCommerce
Machine Learning Applications
Transport
SAFETY MONITORING, AIR TRAFFIC CONTROL ETC.
Healthcare
Financial Services
DRUG DISCOVERY, DISEASE DIAGNOSIS, ROBOTIC SURGERY
ALGORITHMIC TRADING, PORTFOLIO MANAGEMENT, FRAUD DETECTION

Mission and Characteristics of AI

- your idea?

Mission and Characteristics of AI

Mission:

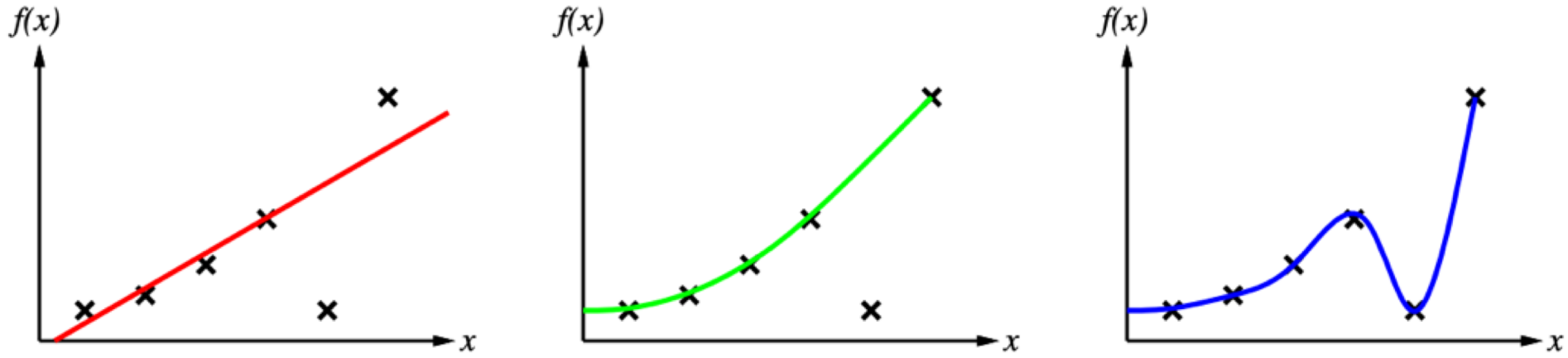
- Problem-solving and decision-making.
- Solve complex problems more efficiently.

Characteristics:

- Learning ability (prediction)
- Optimal solution

Objective of Learning

- Learn to generalize from a finite set of examples
- The learnt function then can predict output y given a new input $\mathbf{x}$



The General Model of Learning from Examples

- Suppose that there is a functional relationship between two sets of objects X and Y:

$$f: X \rightarrow Y$$

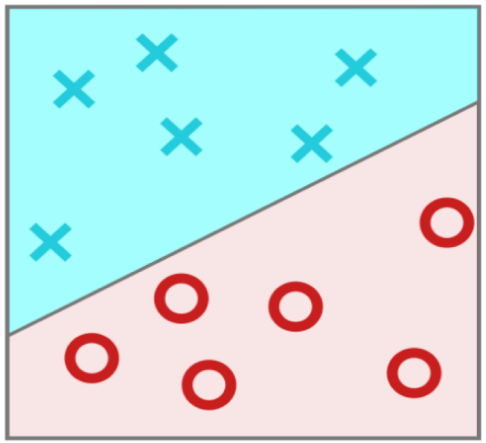
- Given a finite set of examples:

$$D = \{(\mathbf{x}_i, y_i) \mid i=1,2,\ldots,N\} \text{ , where } \mathbf{x}_i \in X \text{ and } y_i \in Y$$

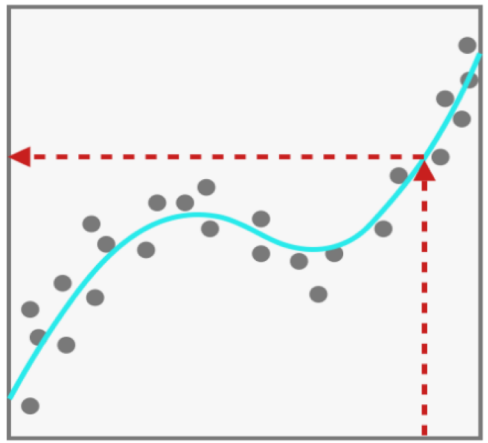
- The task here is to derive (i.e. to learn) the objective function f

Classification and Regression

- $y = f(\mathbf{x})$
- If y is the real value ie Y = R then we have a regression problem
- If y is a value in a given finite discrete set, then we have a classification problem

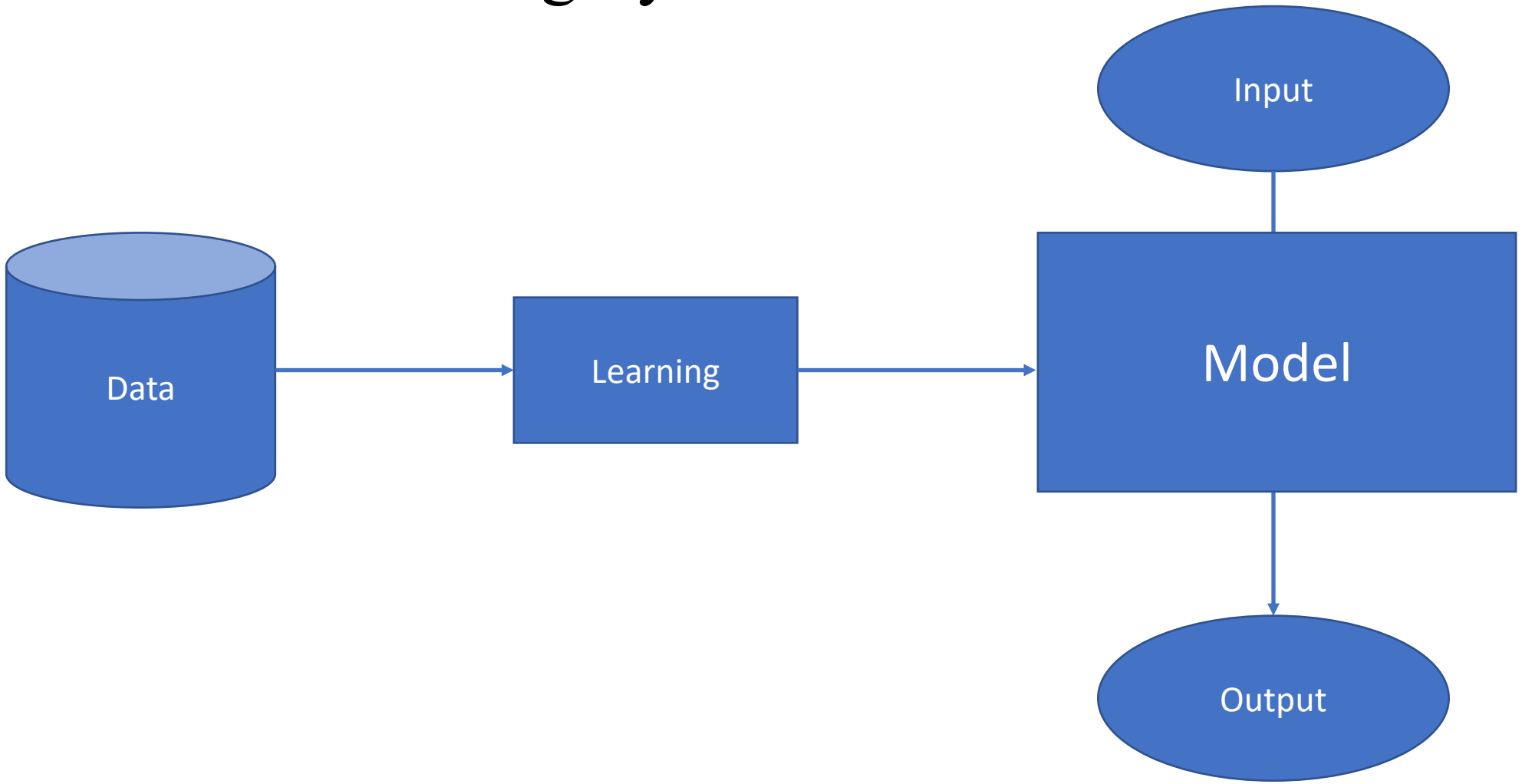


Classification



Regression

A General Paradigm of Machine Learning Systems



Data Representation

- **x** is a vector of features

$$\mathbf{x} = (x_1, x_2, ..., x_d)$$

$$\mathbf{X} = R^d$$

- y is a real number in the regression problem
- y is in classification problem:
  - binary classification, y = {0,1} or {-1,+1}
  - multiple classes: y = {1, 2, ..., k} or one-hot vector (0,...0,1,0,...0)

Loss function

- Suppose that ($\boldsymbol{x}$,y) is an example. We want to find the difference between the ground true value y and the predicted value h($\boldsymbol{x}$)
- For regression:

$$L(y,h(x)) = \left(y - h(x)\right)^2$$

- For classification:

$$L(y,h(x)) = \begin{cases} 0 \text{ nếu } y = h(x) \\ \\ 1 \text{ nếu } y \neq h(x) \end{cases}$$

Expected Risk and Empirical Risk

- Expected risk/loss is the mean of L(y,h(**x**)) over the whole space ***X*** x ***Y***

$$R(h) = \iint L(y, h(x))p(x, y)dxdy$$

- Empirical risk/loss is the mean of L(y,h(**x**)) over the training dataset D

$$R_{emp}(h) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, h(x_i))$$

Empirical Risk

- For regression:

$$R_{emp}(h) = \frac{1}{N}\sum_{i=1}^{N}(y_i - h(x_i))^2$$

- For classification:

$$R_{emp}(h) = \frac{1}{N}\sum_{i=1}^{N}\delta(y_i, h(x_i))$$

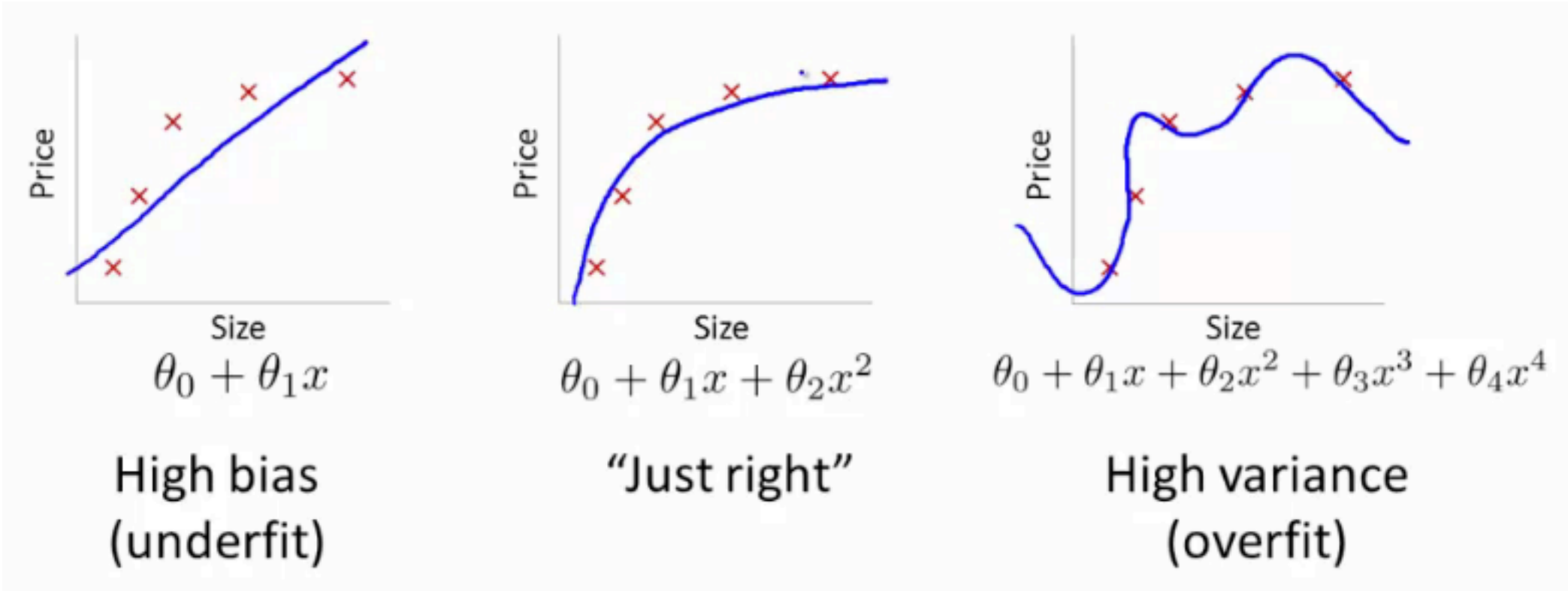
Empirical risk minimization inductive principle. (Nguyên lý quy nạp cực tiểu sai số thực nghiệm)

- We will consider the objective function f by the approximation function g as follows:

$$\mathbf{H} = \{h : X \to Y\}$$

$$g = \arg\min_{h} R_{emp}(h)$$

Overfitting



What are ML components in AI systems?

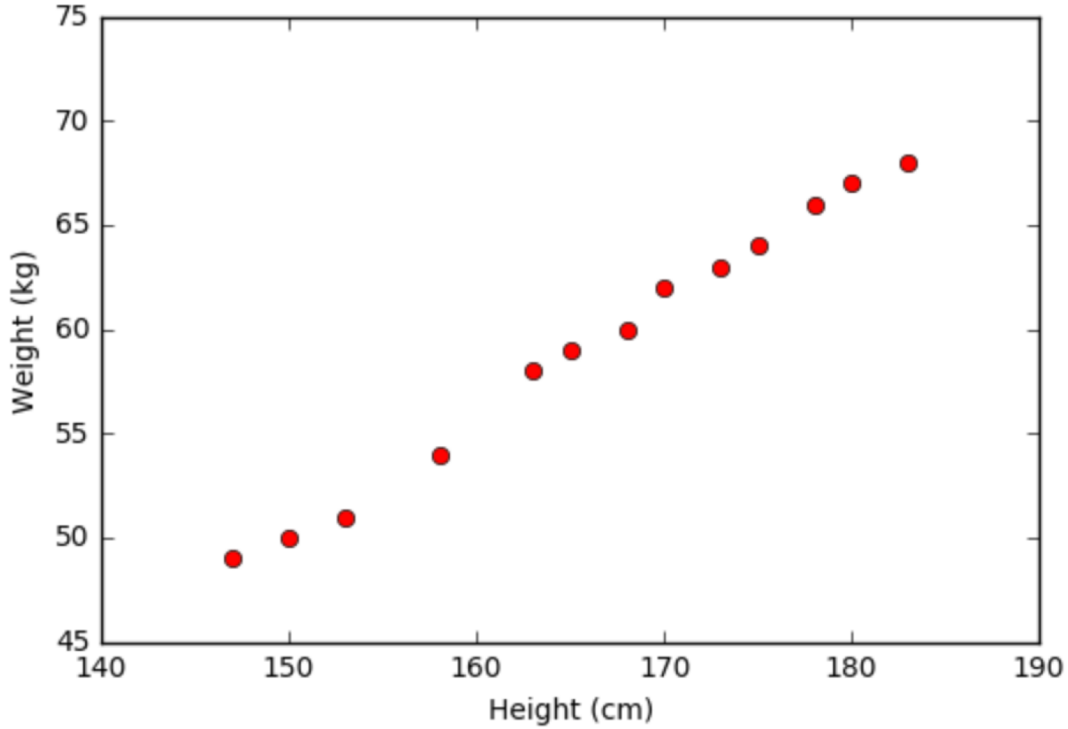
- Model
- Learning algorithms/methods

Outline

1. Machine learning and Artificial Intelligence
2. A general paradigm for AI systems and the ML component
3. Examples of ML models
4. Categorization of Machine learning problems and methods

ML model: example 1

Chiều cao (cm)	Cân nặng (kg)	Chiều cao (cm)	Cân nặng (kg)
147	49	168	60
150	50	170	72
153	51	173	63
155	52	175	64
158	54	178	66
160	56	180	67
163	58	183	68
165	59		



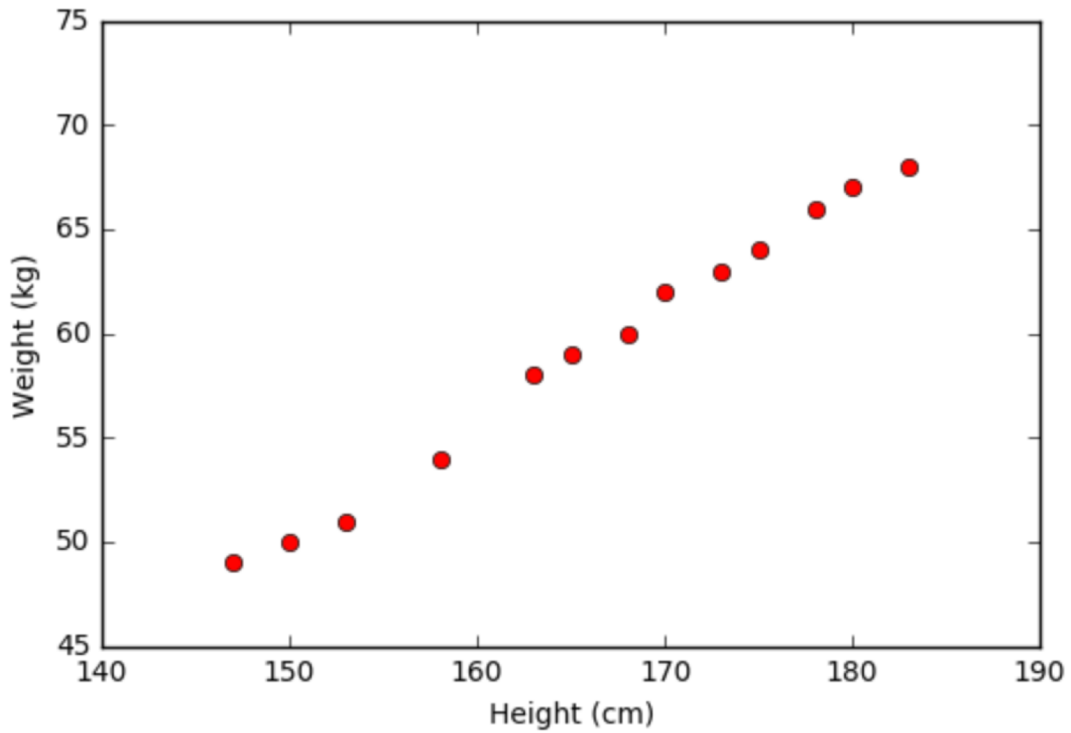
(cân nặng) = w_1*(chiều cao) + w_0

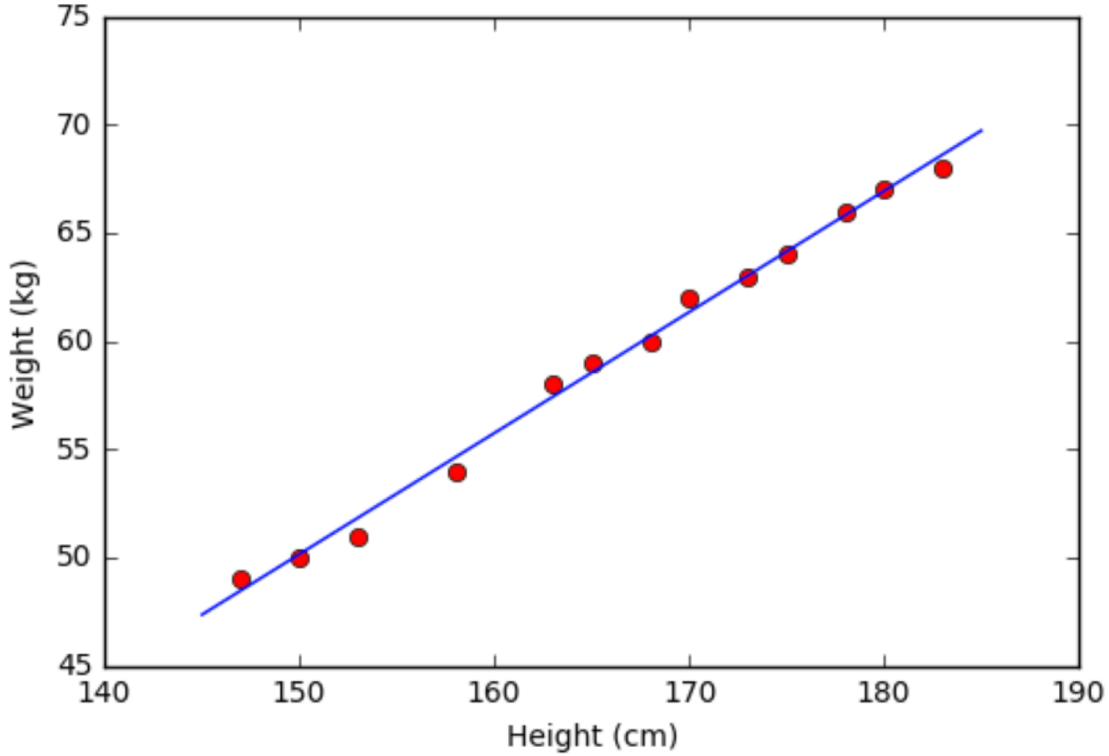
$y = ax + b$

(cân nặng) = w_1*(chiều cao) + w_0

```
w =  [[-33.73541021]
 [  0.55920496]]
```

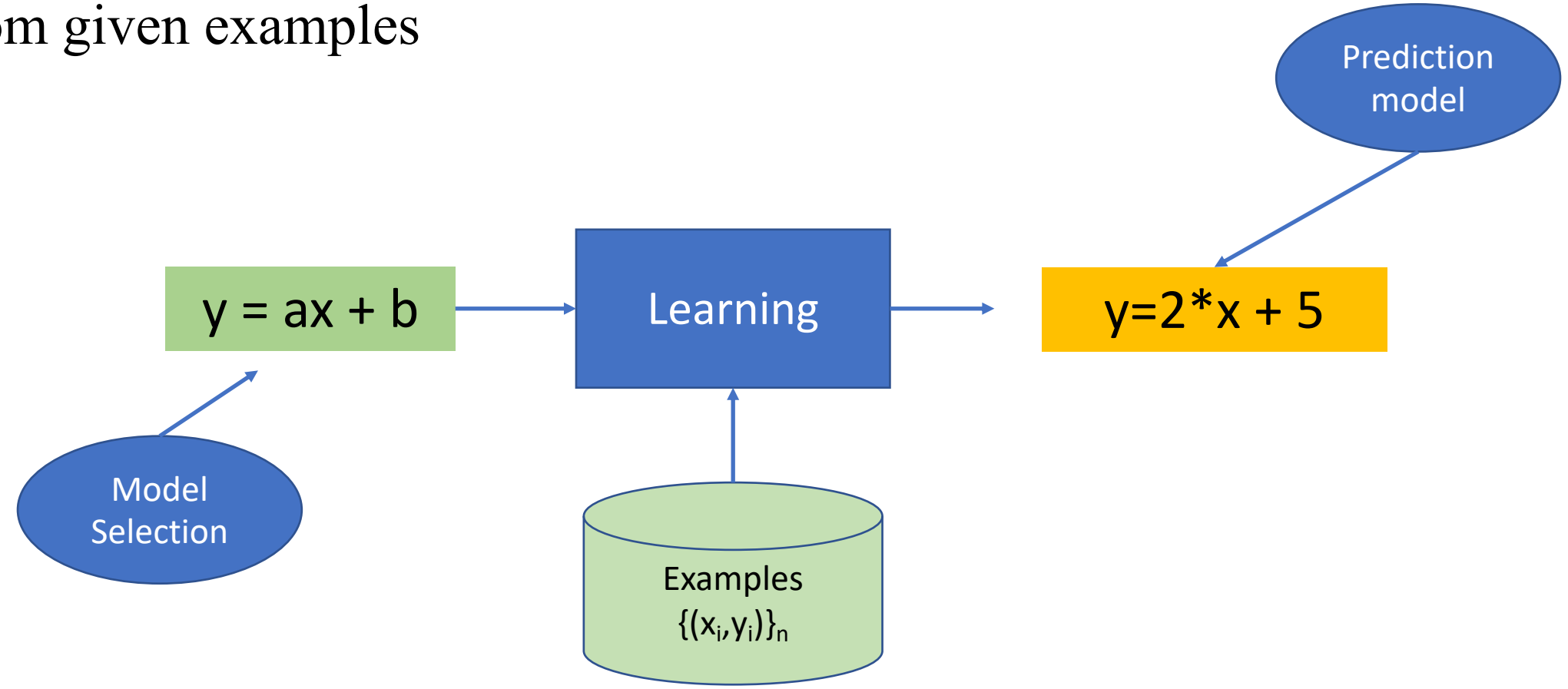
$$y = ax + b$$





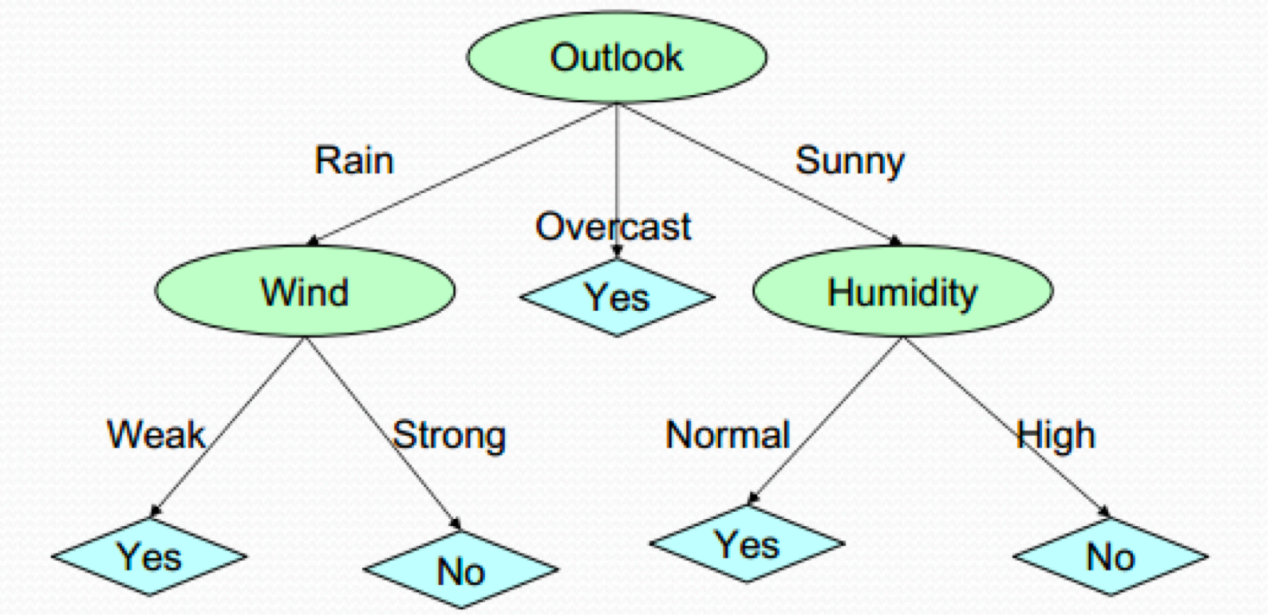
How to understand Learning?

=> learn values for model's parameters by inferencing from given examples

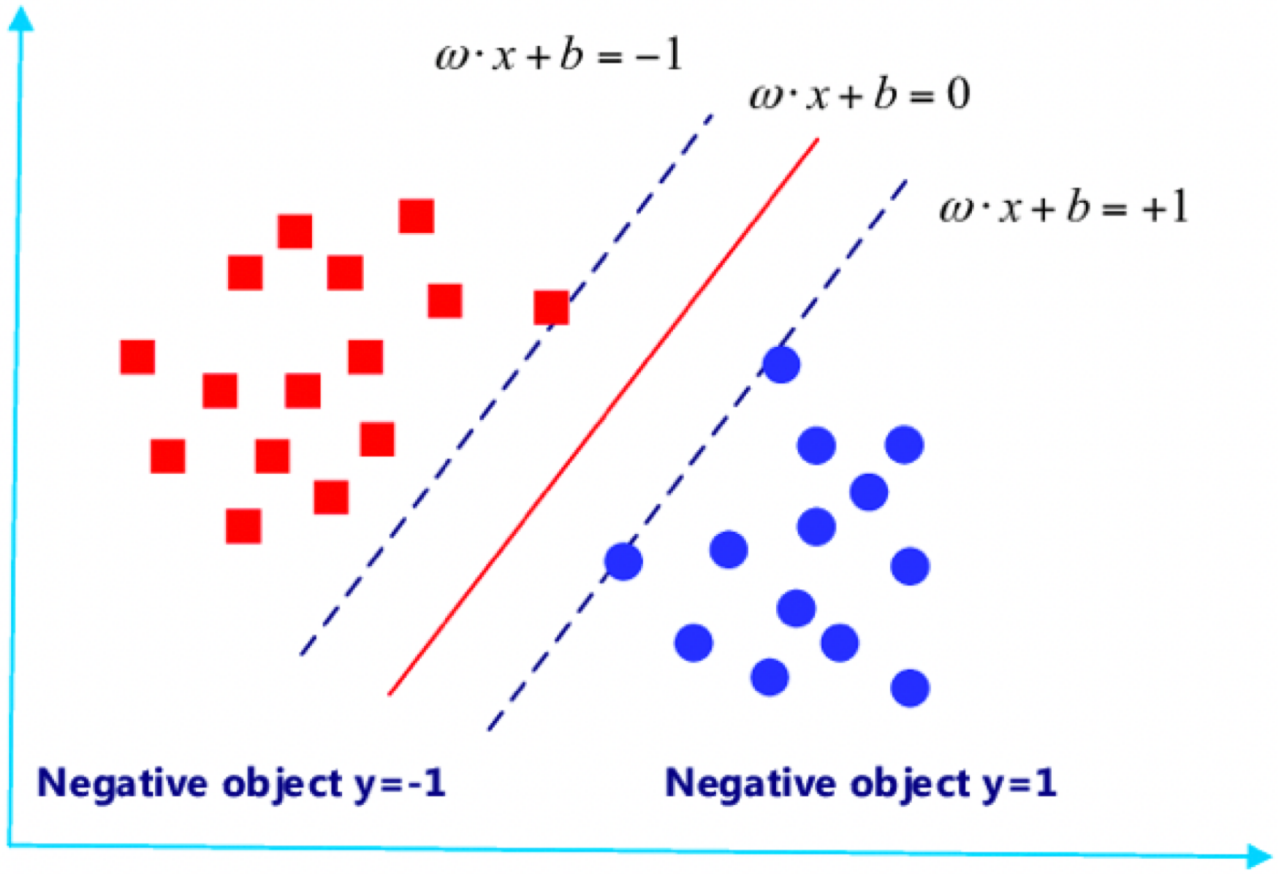


ML model: example 2

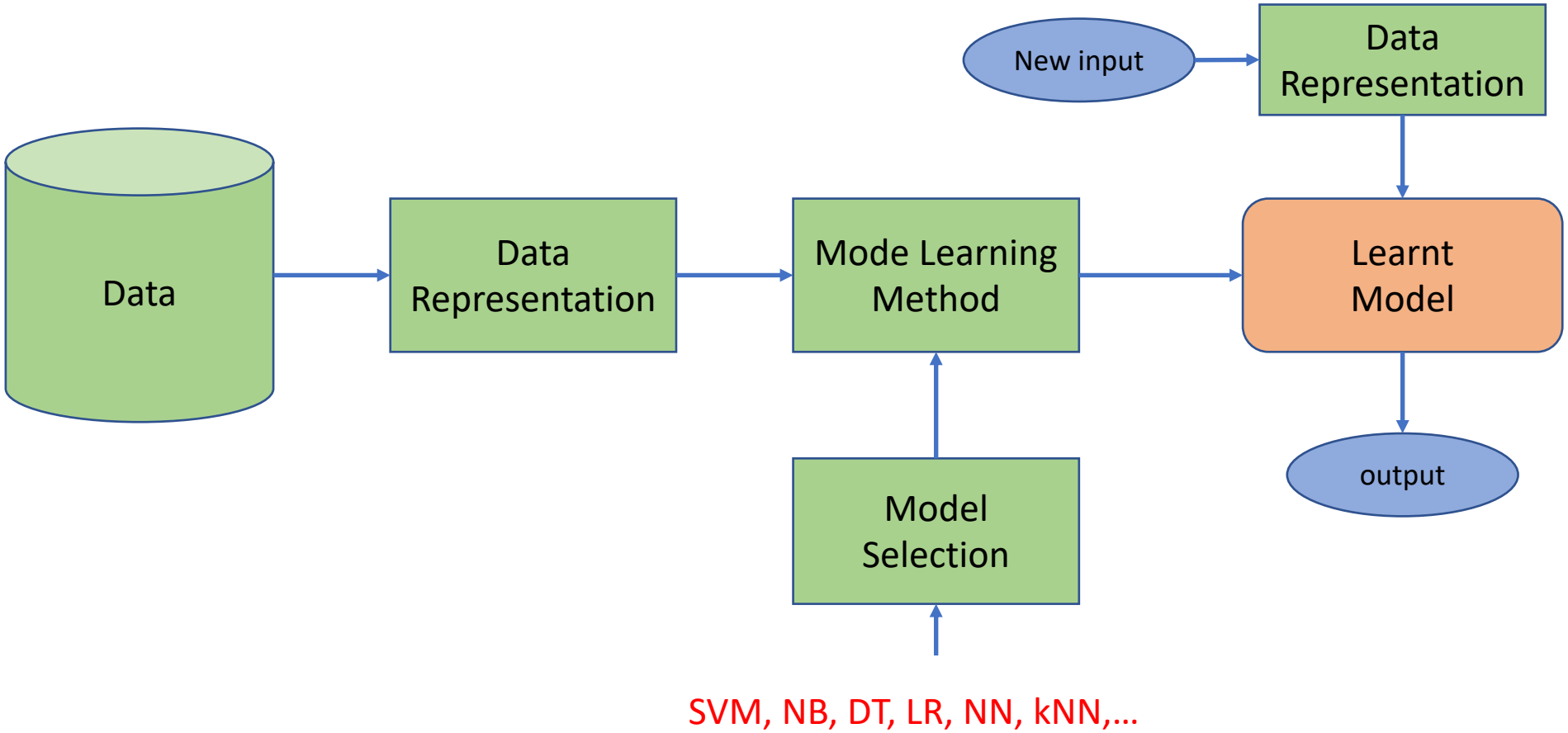
Outlook	Tempreature	Humidity	Windy	Class
sunny	hot	high	false	N
sunny	hot	high	true	N
overcast	hot	high	false	P
rain	mild	high	false	P
rain	cool	normal	false	P
rain	cool	normal	true	N
overcast	cool	normal	true	P
sunny	mild	high	false	N
sunny	cool	normal	false	P
rain	mild	normal	false	P
sunny	mild	normal	true	P
overcast	mild	high	true	P
overcast	hot	normal	false	P
rain	mild	high	true	N



ML model: example 3



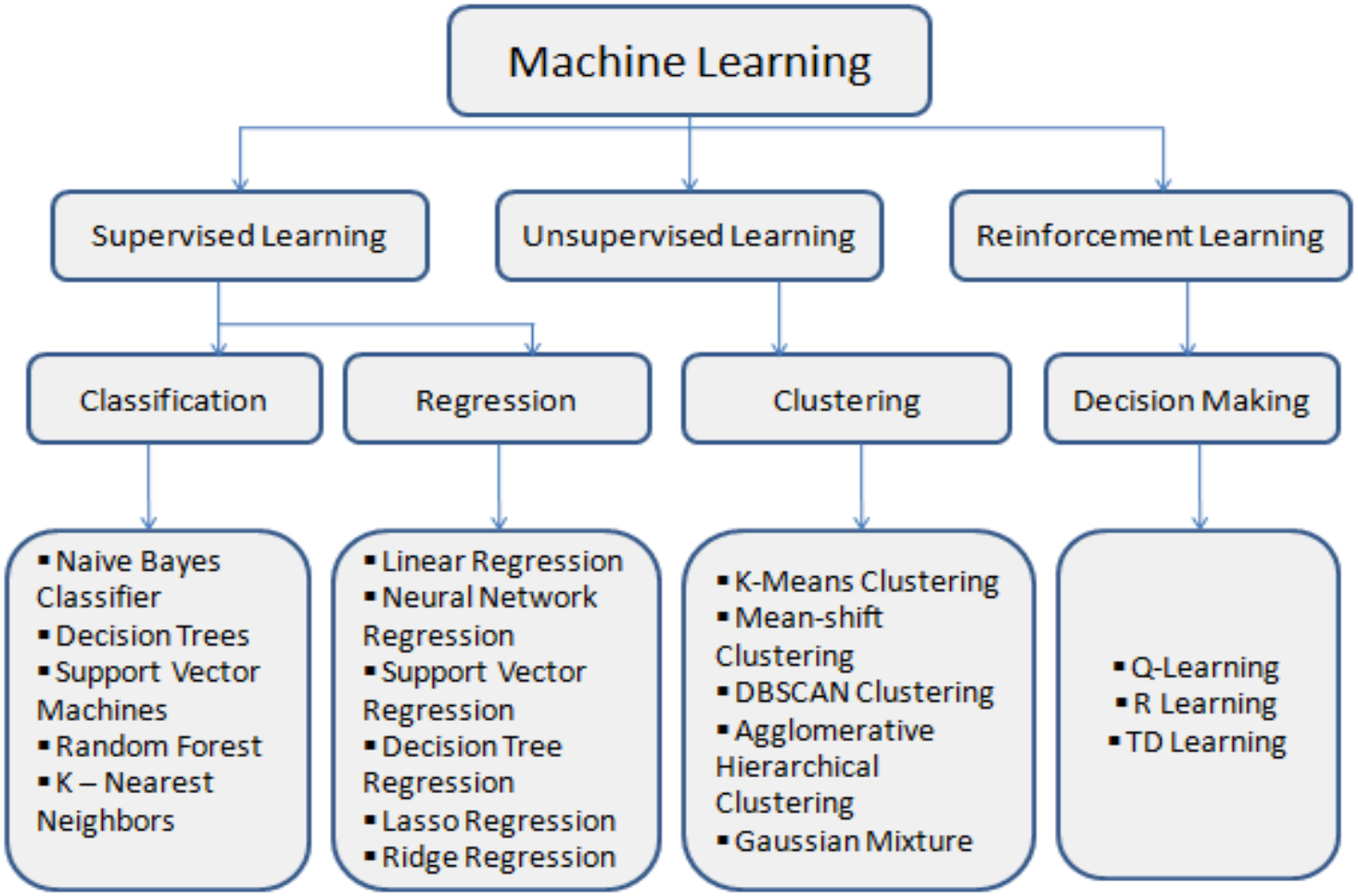
A General Machine Learning Diagram



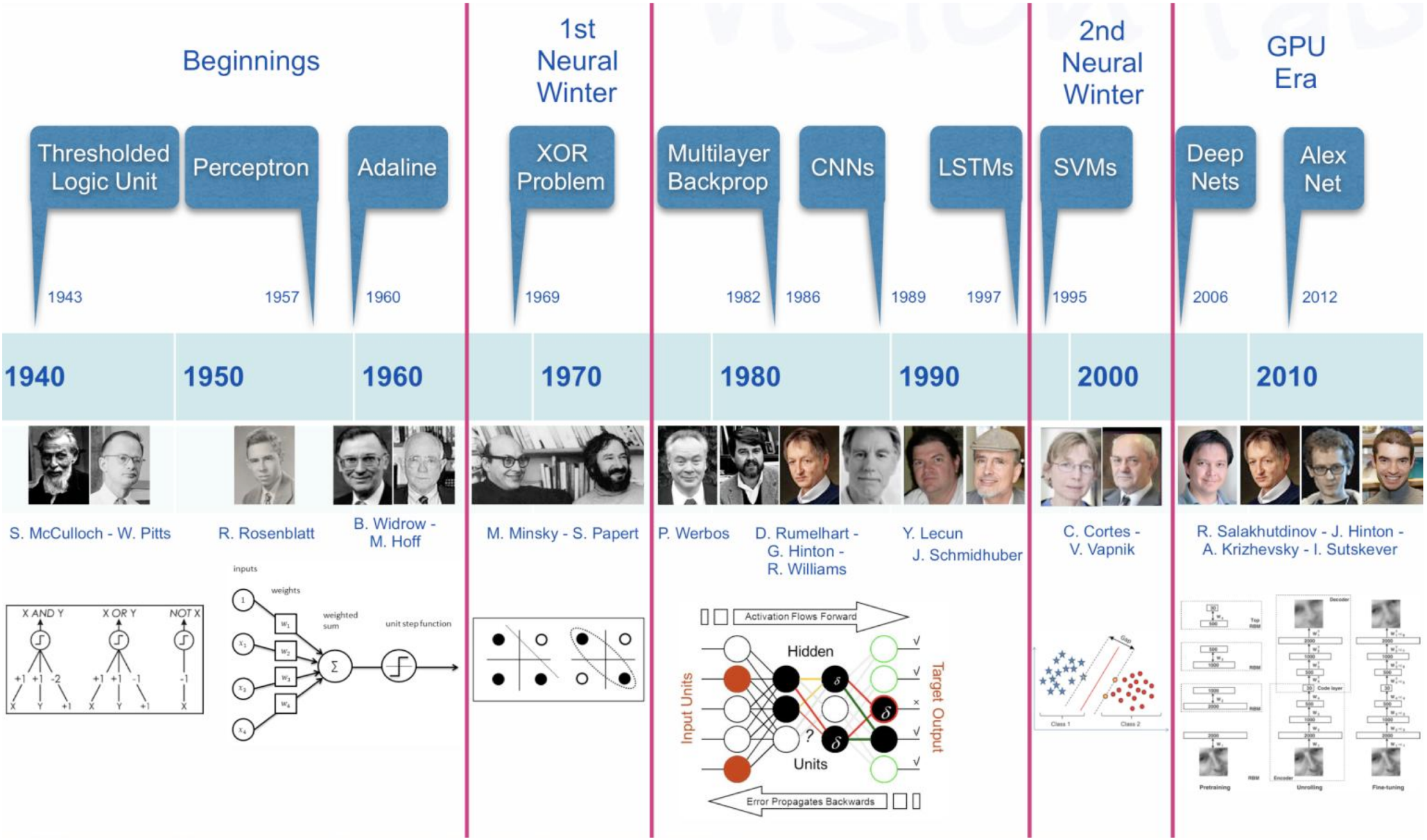
What should be learnt in Machine Learning?

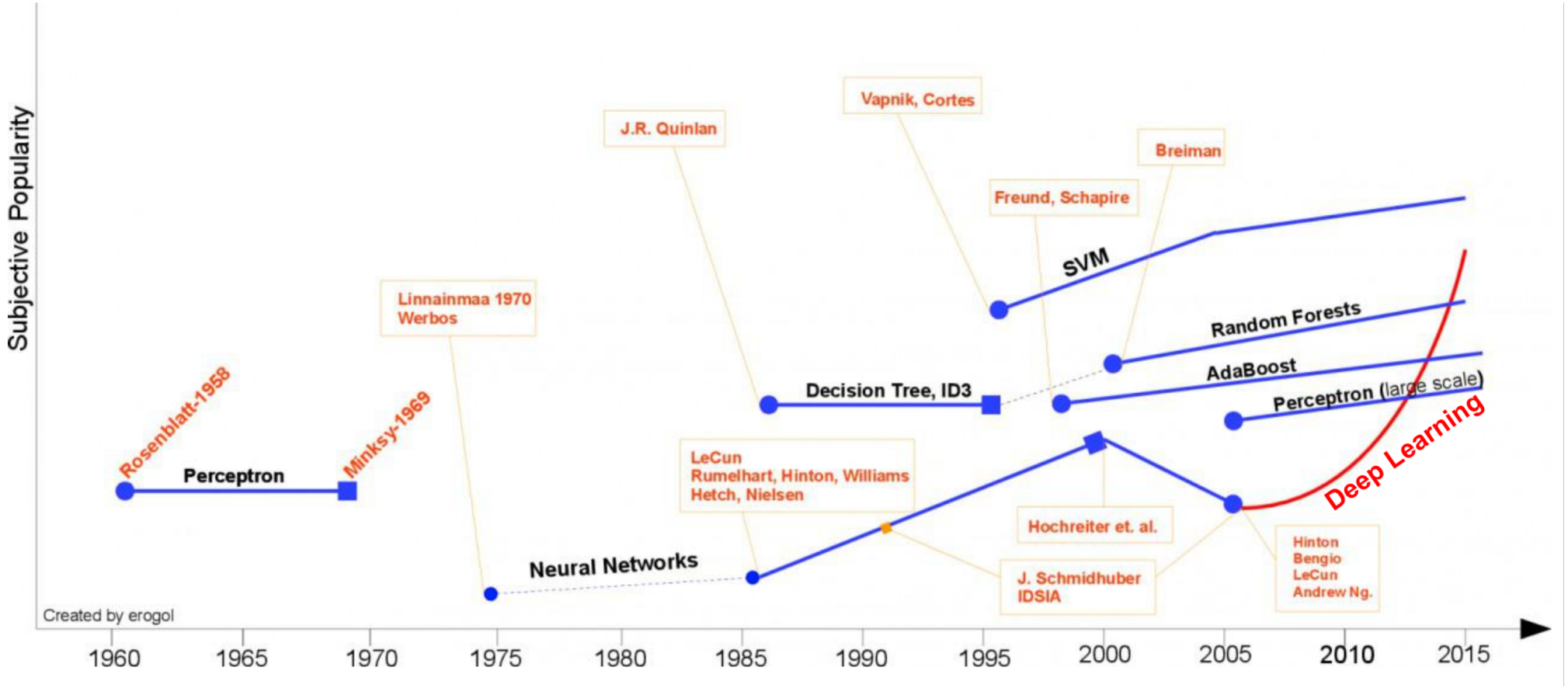
- Understanding ML models
  - Architecture, Structure, or form of the models
  - Computation, inference/reasoning
- Methods/Algorithms for learning models' parameters
- Model evaluation
- Data Processing: data collection; feature selection; dimensional reduction, noisy filtering,…
- Other issues: overfitting,

Types of Machine Learning



History of ML




Subjective Popularity
Vapnik, Cortes
J.R. Quinlan
Breiman
Freund, Schapire
SVM
Linnainmaa 1970
Werbos
Random Forests
AdaBoost
Perceptron (large scale)
Rosenblatt-1958
Minksy-1969
Decision Tree, ID3
Perceptron
LeCun
Rumelhart, Hinton, Williams
Hetch, Nielsen
Deep Learning
Hochreiter et. al.
Hinton
Bengio
LeCun
Andrew Ng.
J. Schmidhuber
IDSIA
Neural Networks
Created by erogol
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015

Objective of this course?

- Understand:
  - ML methods
  - ML issues
- Applying:
  - Applying ML for real prediction
- Analyze and Evaluation
  - Compare and evaluate the strengths and weaknesses of the methods
- Building industrial ML systems

Learning method

- Active learning
- Project based

Course Assessment

- Exercises (small projects): 40%
- Theorical self-study (books): 20%
- Final Project: 40%