Machine Learning vs. Deep Learning 概念

林郁修

Yu-Hsiu Lin
yhlin@ntut.edu.tw
https://sites.google.com/view/aici

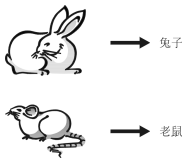
Graduate Institute of Automation Technology,
National Taipei University of Technology

Outline

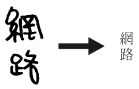
- 可解決的問題
- 機器學習的架構
- 學習方法的類型
- 實踐學習方法的流程

可解決的問題 (examples of machine learning)

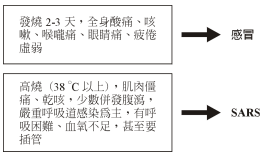
1. 辨識/分類
(pattern recognition/classification)



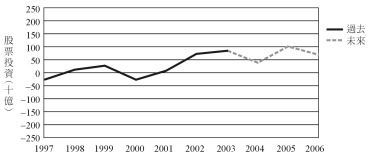
2. 歸納推演
(generalization)



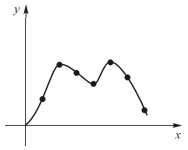
3. 評估、診斷/決策
(diagnosis/decision)



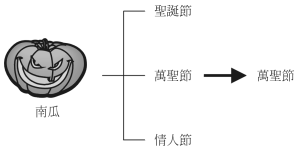
4. 預測
(predictions)



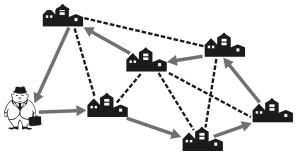
5. 近似
(function approximation)



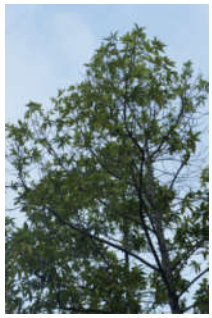
6. 聯想
(association)



7. 最佳化
(optimization)

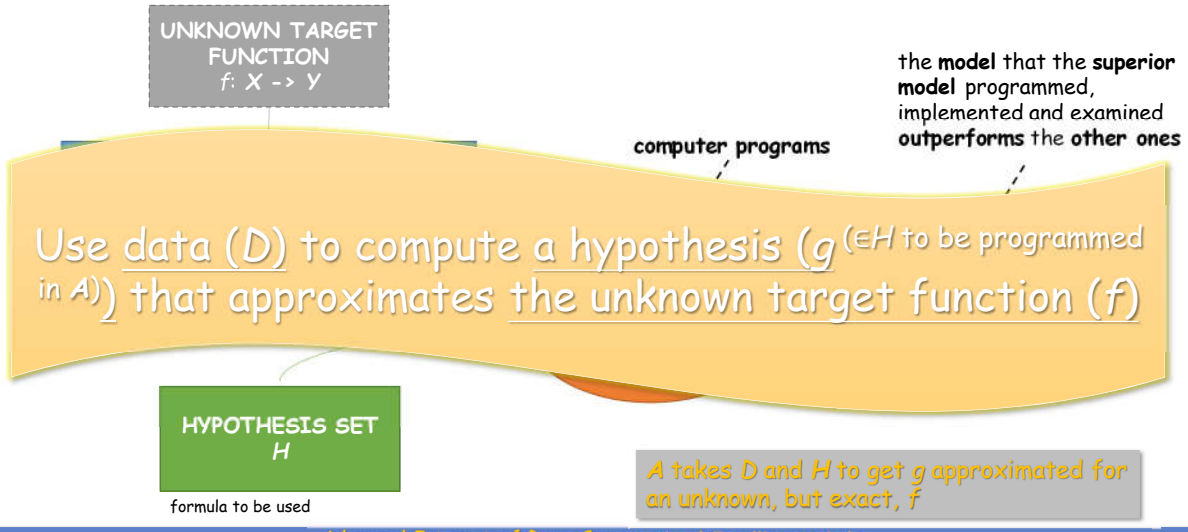


- 適合使用**機器學習**解決的問題通常有以下的特徵
  - 待解問題及其相關條件難以完整*f*定義
  - 待解問題非常複雜或是屬非線性的問題--**無法**由一連串已知的**數學方程式/*f*來描述**並求得解答者
  - …

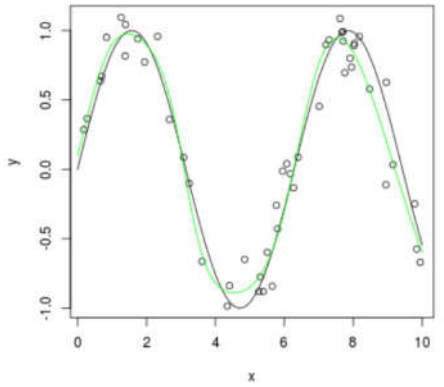


- Define trees and Hand-program: **difficult**
- Learn from data/observations and Recognize: a **3-year-old can do so**

機器學習的架構 (formulation of learning)

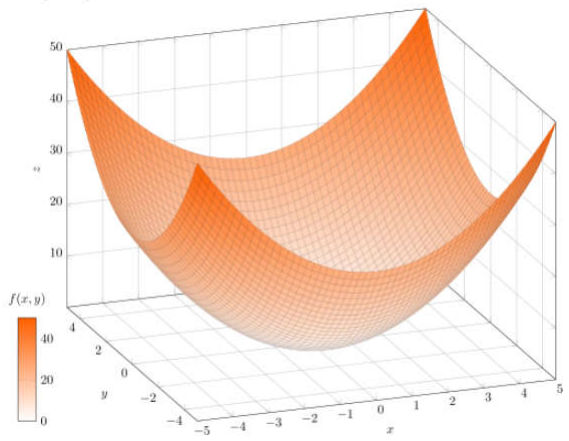


A **function approximation** problem:
$y=f(x)=\sin(x)$



true
approximated (model)
noisy observations

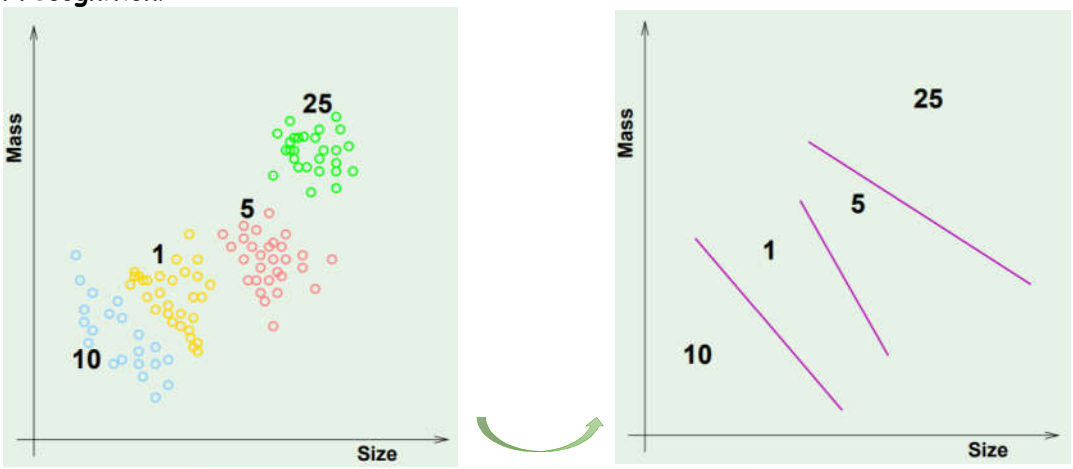
$z = f(x, y) = x^2 + y^2$



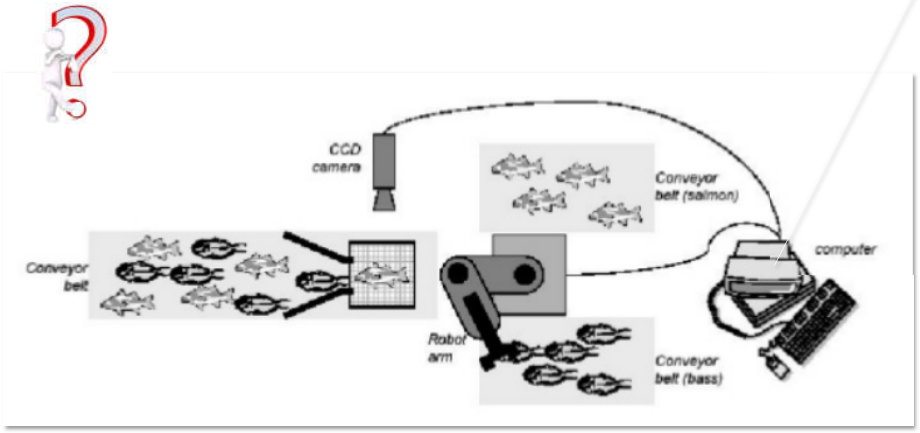
Euclidean n-space:

$\text{Dist}(\boldsymbol{x}, \boldsymbol{y})=\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$

- One of the simplest hypotheses, ***perceptron(s)***, for solving a classification problem--**coin recognition**!

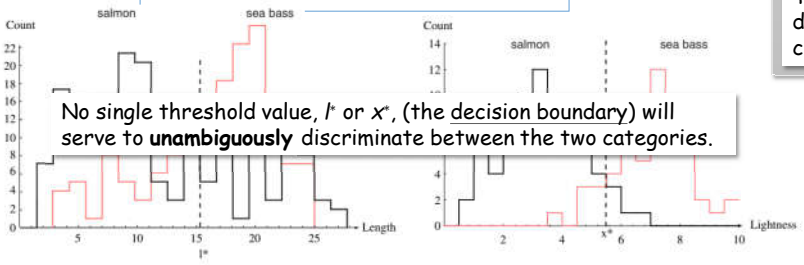


- Another example

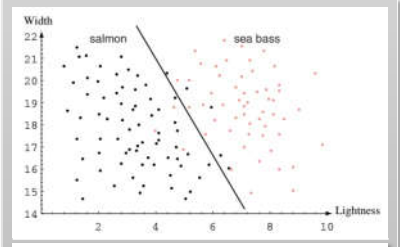


- 魚罐頭工廠--**鮭魚**、**鱸魚**分類
  - Digital Image Processing
    - ✓ Feature Extraction
  - AI/ML
  - …

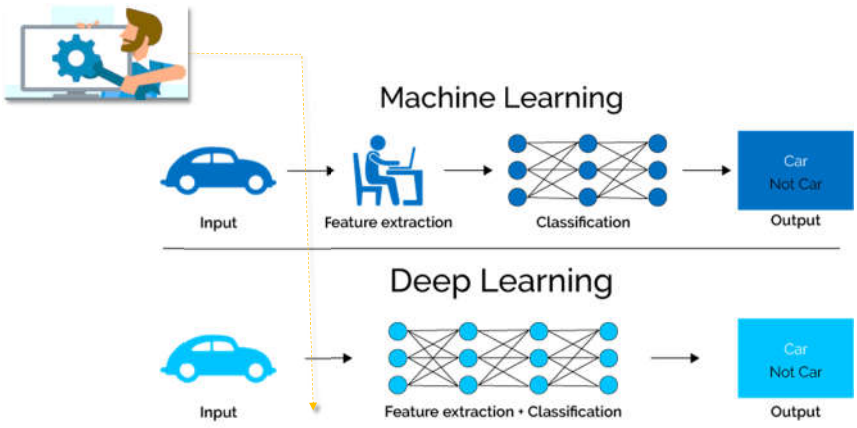
https://www.youtube.com/watch?v=Mh2mZaqn8Vk
https://www.youtube.com/watch?v=3t2A27ppWMI



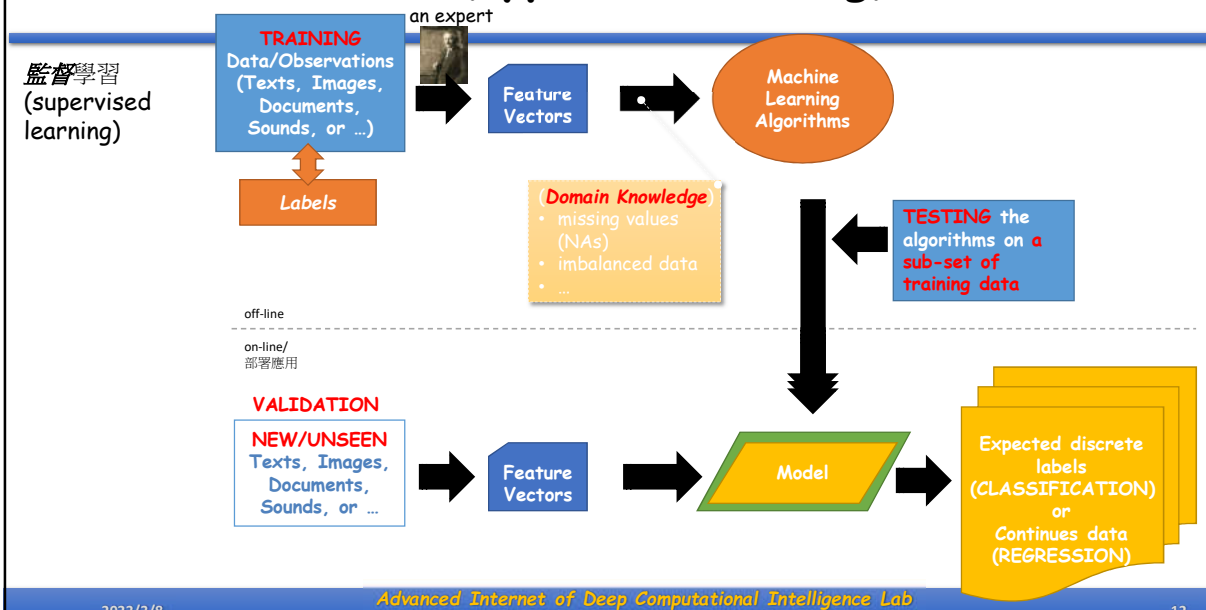
No single threshold value, l^* or x^*, (the decision boundary) will serve to **unambiguously** discriminate between the two categories.



The dark line might serve as a decision boundary, *g*, of your classifier, *A*.

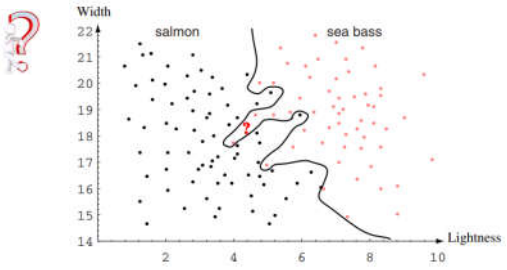


學習方法的類型 (types of learning)

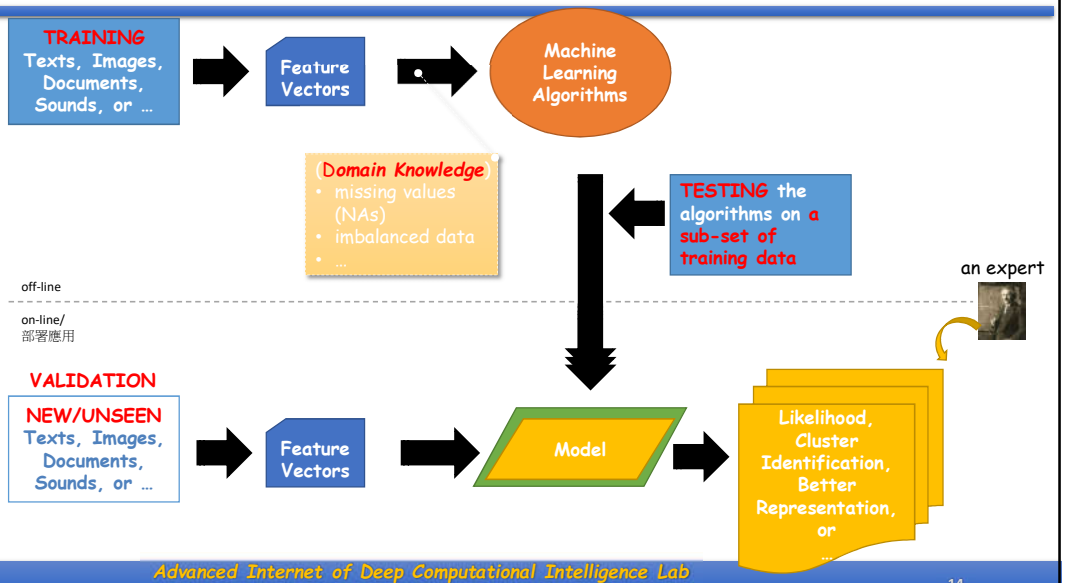


- 藉由訓練(training/learning from data)模型的過程，降低模型的輸出值與目標值 (ground truth: 現實生活中的實際值) 之間的誤差值 (差異)

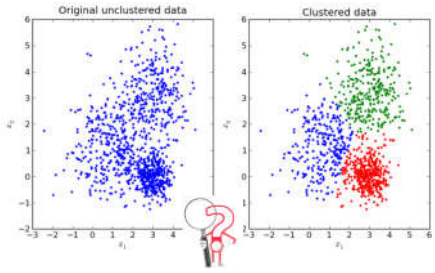
  - 過度學習 (over-fitting)

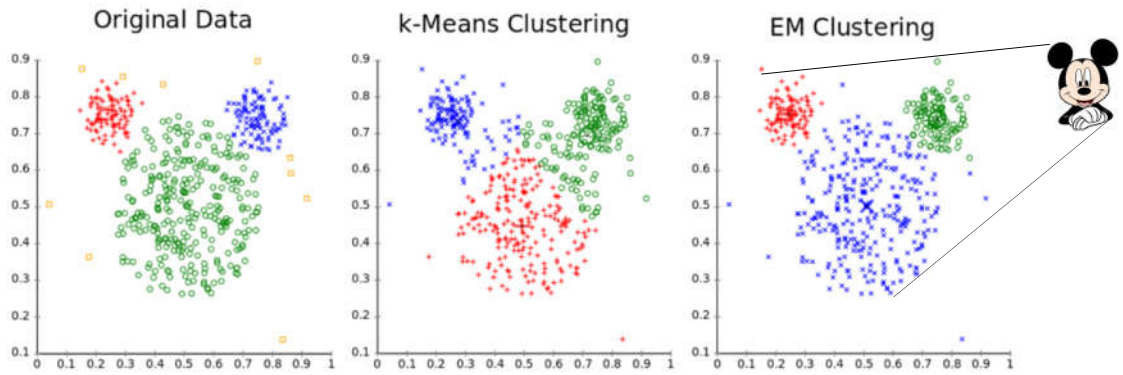


非監督學習 (un-supervised learning)

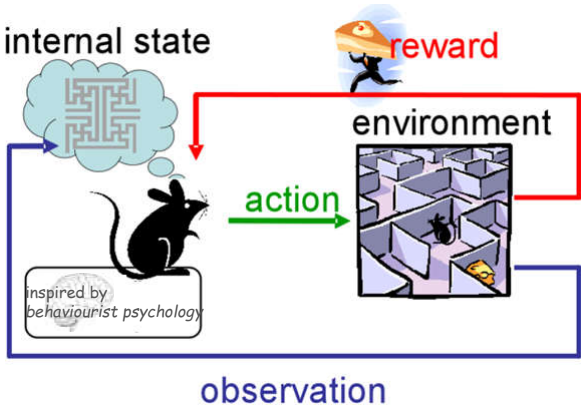


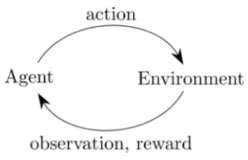
- 無目標值
- 適合用於解決*聚類*的問題
  - 獲取不同資料群組之間的屬性/關聯關係





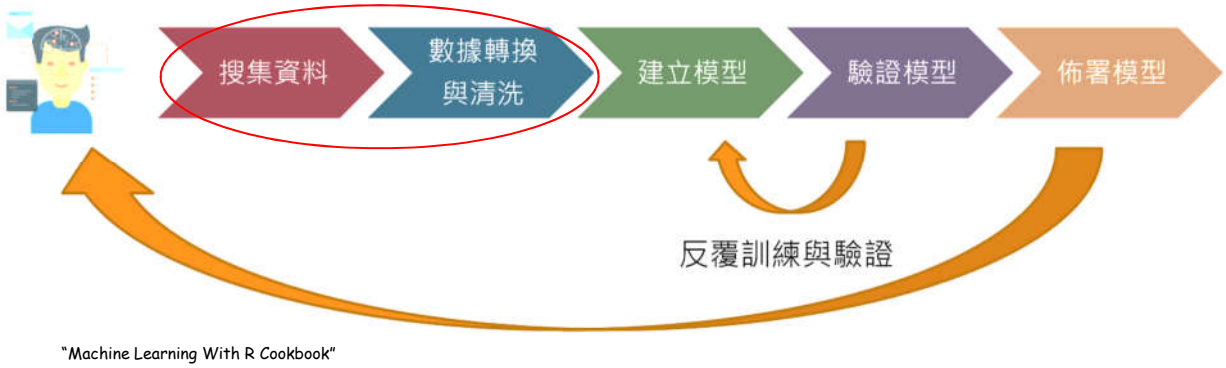
增強學習
(Reinforcement Learning)





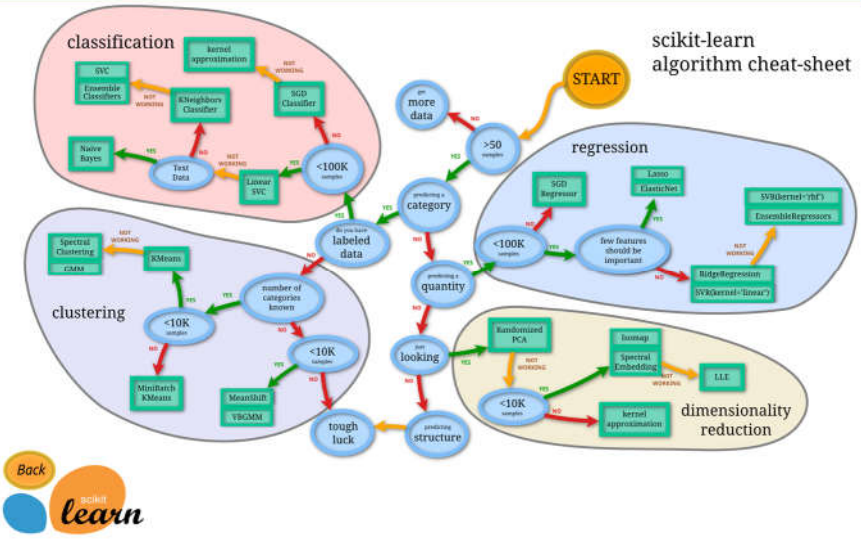
實踐學習方法的流程 (Learning Process)

- 蒐集資料 (鑑往知來)
  - Labeled/Un-labeled
- 萃取特徵
  - 以分類問題為例，*特徵工程*於可能區分各類別的*屬性*
- 採取機器學習演算法
  - 決定演算法
  - 以隨機比例抽樣，將整體資料集分成：訓練集(60%)、測試集(20%)及驗證集(20%)
  - 訓練、測試及驗證(避免over-fitting)演算法於真實世界的問題解決
- 應用所訓練完成之模型於真實世界的問題



"Machine Learning With R Cookbook"

https://scikit-learn.org/stable/



- *The End*