

網研PYTHON

WEEK6 11/20



常常聽到人家說機器學習、
深度學習、監督、非監督，
常深是不是聽到很混亂呢？



ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



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.

source: <https://blogs.nvidia.com/blog/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>



(圖片來源：OmniXRI 整理繪製)

人工智慧(Artificial Intelligence)

機器學習(Machine Learning)

監督學習
(Supervised)

分類 迴歸

類神經
網路(NN)

非監督學習
(Unsupervised)

聚類

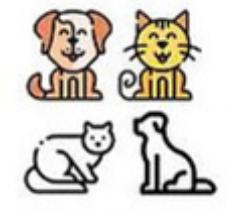
增強學習
(Reinforcement)

分類

深度學習(Deep Learning)
DNN CNN RNN GAN



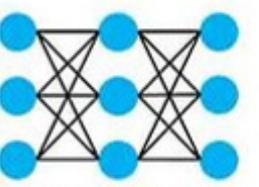
Machine Learning



Data Input



Feature Extraction



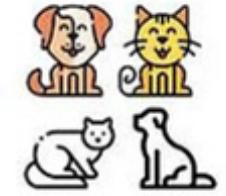
Classifier

Cat
Dog

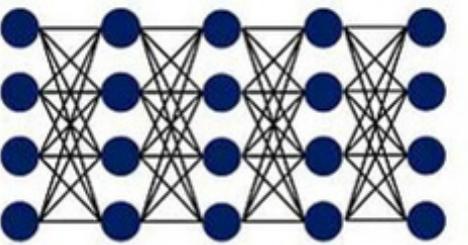
Output

先比較
機器&深度

Deep Learning



Data Input

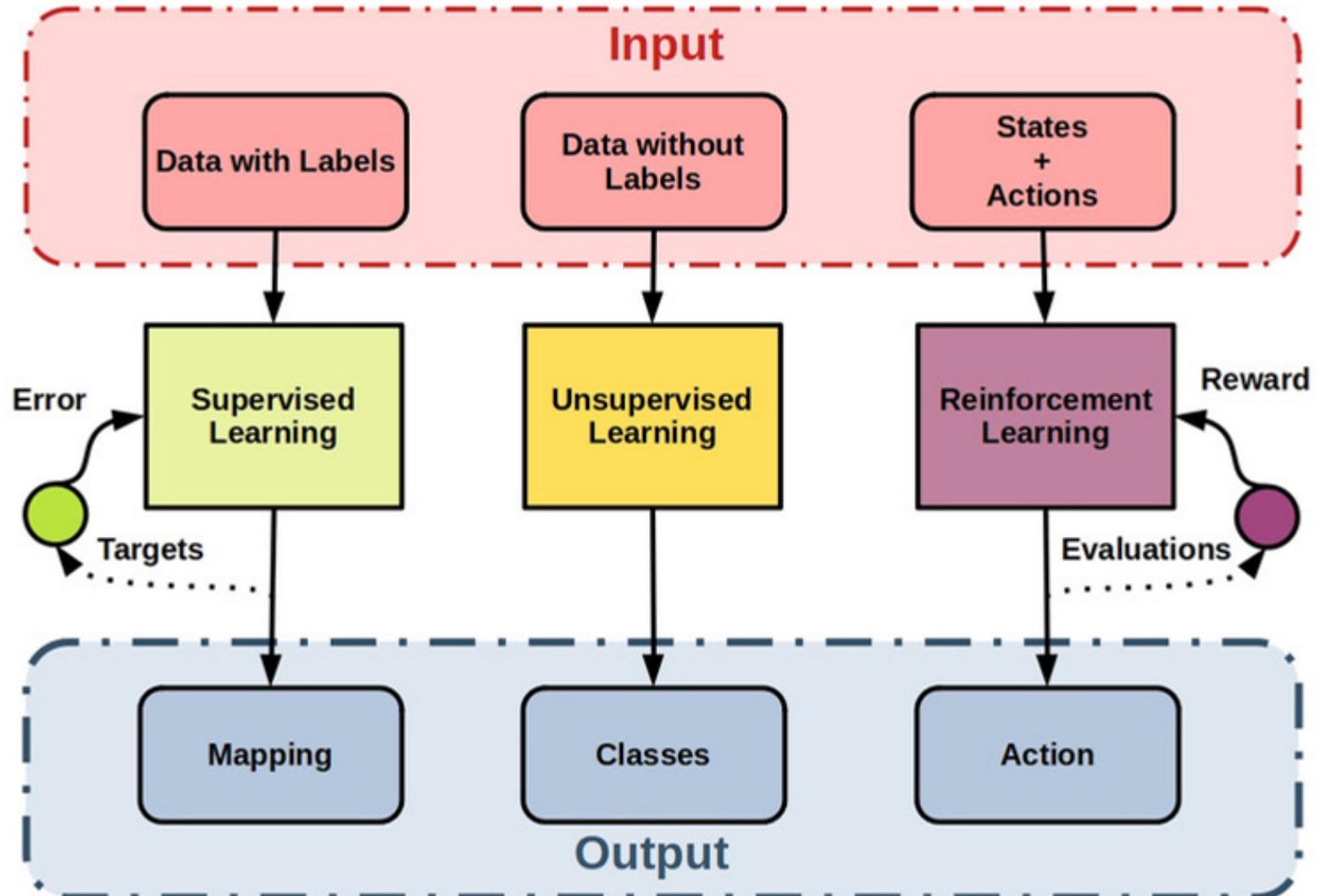
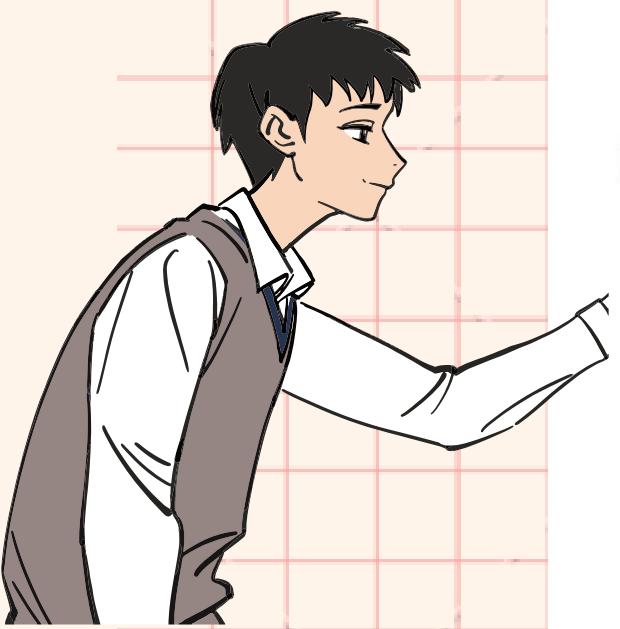


Feature Extraction + Classifier

Cat
Dog

Output





- <https://starship-knowledge.com/supervised-vs-unsupervised-vs-reinforcement>

整理如下：

類別	功能	演算法
監督式學習 Supervised	預測 Predicting	Linear Regression Decision Tree Random Forest Neural Network Gradient Booting Tree
	分類 Classification	Decision Tree Naive Bayes Logistic Regression Random Forest SVM Neural Network Gradient Booting Tree
非監督式學習 Unsupervised	分群 Clustering	K-means
	關聯 Association	Apriori
	降維 Dimension Reduction	PCA





Train/Test

70% for training, and 30% for testing.



Train the model means create the model.

Test the model means test the accuracy of the model.

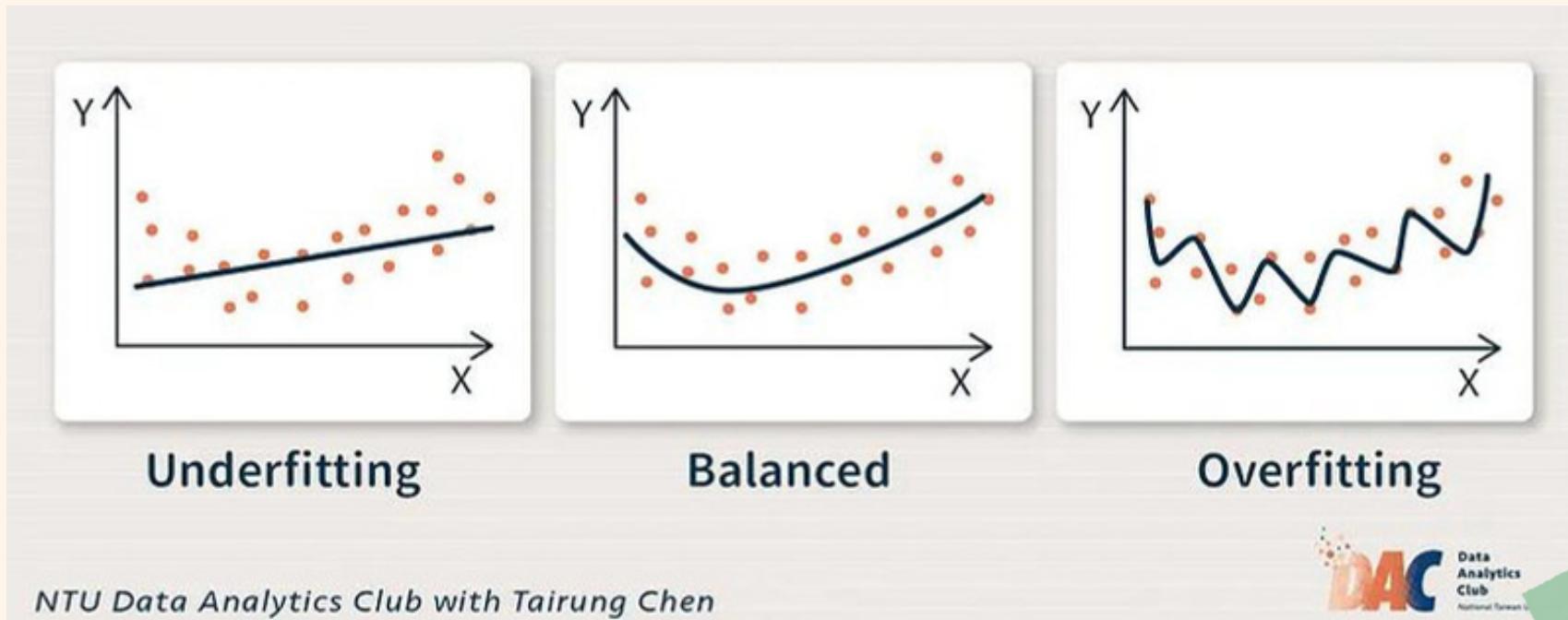
習慣上: 7-3拆資料

理論上應該像是從小資料
預估大世界





Overfitting



不好聽但貼切的比喻: 不要完全變成訓練集的形狀



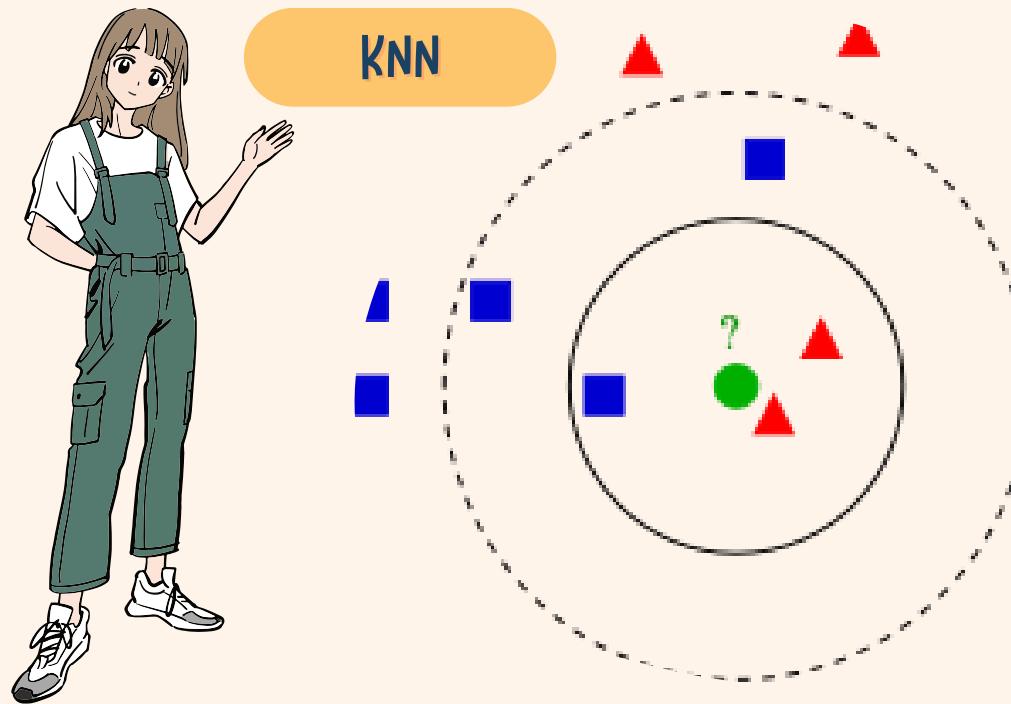
今天來學些算法吧~



<https://iaic.nccu.edu.tw/ai-column/26>

KNN (K Nearest Neighbors)

「近朱者赤，近墨者黑」



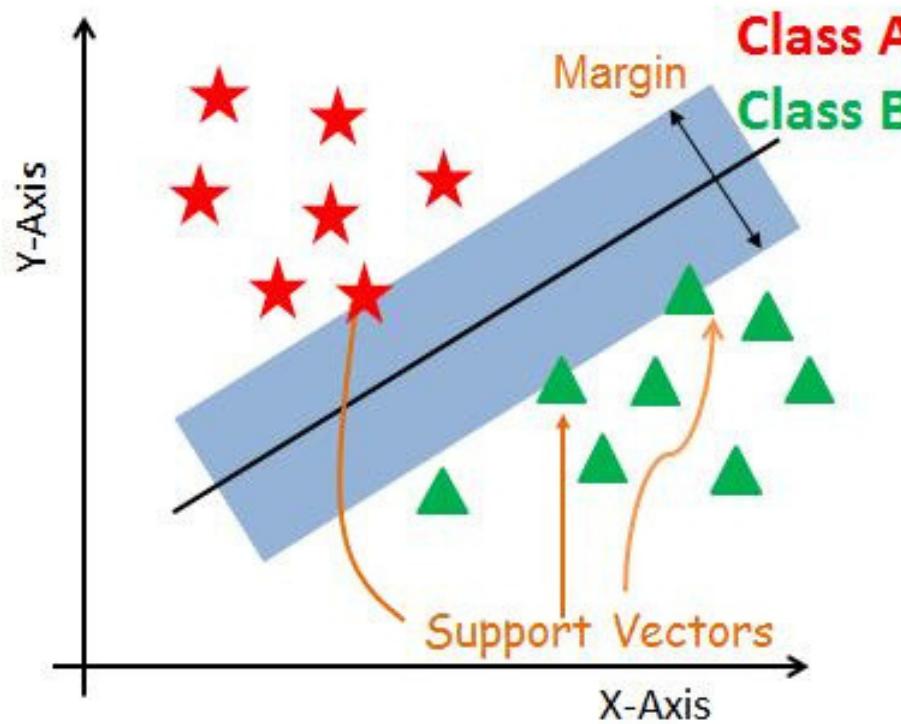
https://pyecontech.com/2020/04/11/python_svm/

<https://medium.com/jameslearningnote/%E8%B3%87%E6%96%99%E5%88%86%E6%9E%90-%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E7%AC%AC3-4%E8%AC%9B-%E6%94%AF%E6%8F%B4%E5%90%91%E9%87%8F%E6%A9%9F-support-vector-machine-%E4%BB%8B%E7%B4%B9-9c6c6925856b>

SVM (Support Vector Machine)

線性、非線性

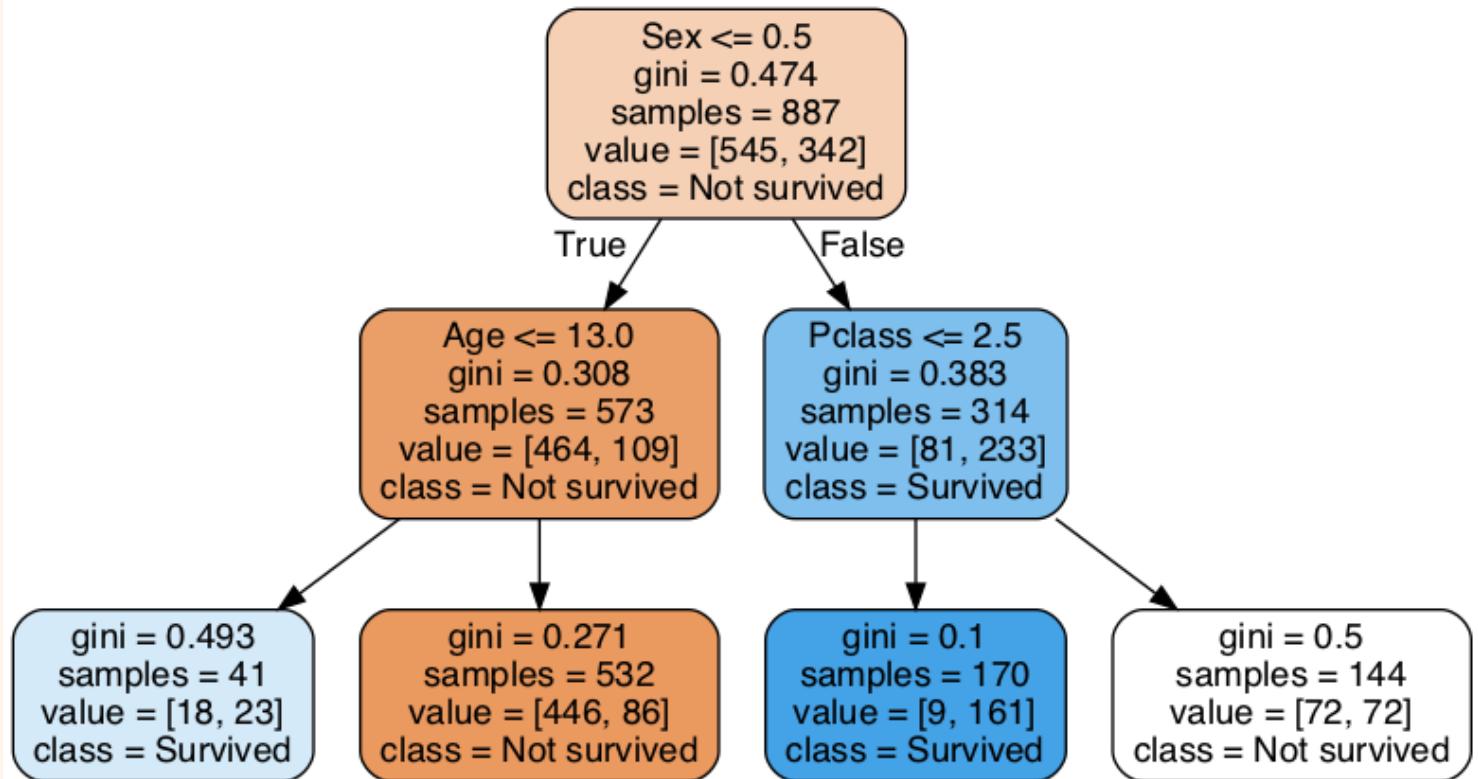
<https://notes.andrewu.tw/2020/%E7%99%BD%E8%A9%B1%E6%96%87%E8%AC%9B%E8%A7%A3%E6%94%AF%E6%8C%81%E5%90%91%E9%87%8F%E6%A9%9F%E4%BA%8C-%E9%9D%9E%E7%B7%9A%E6%80%A7svm/>



SVM



Decision tree



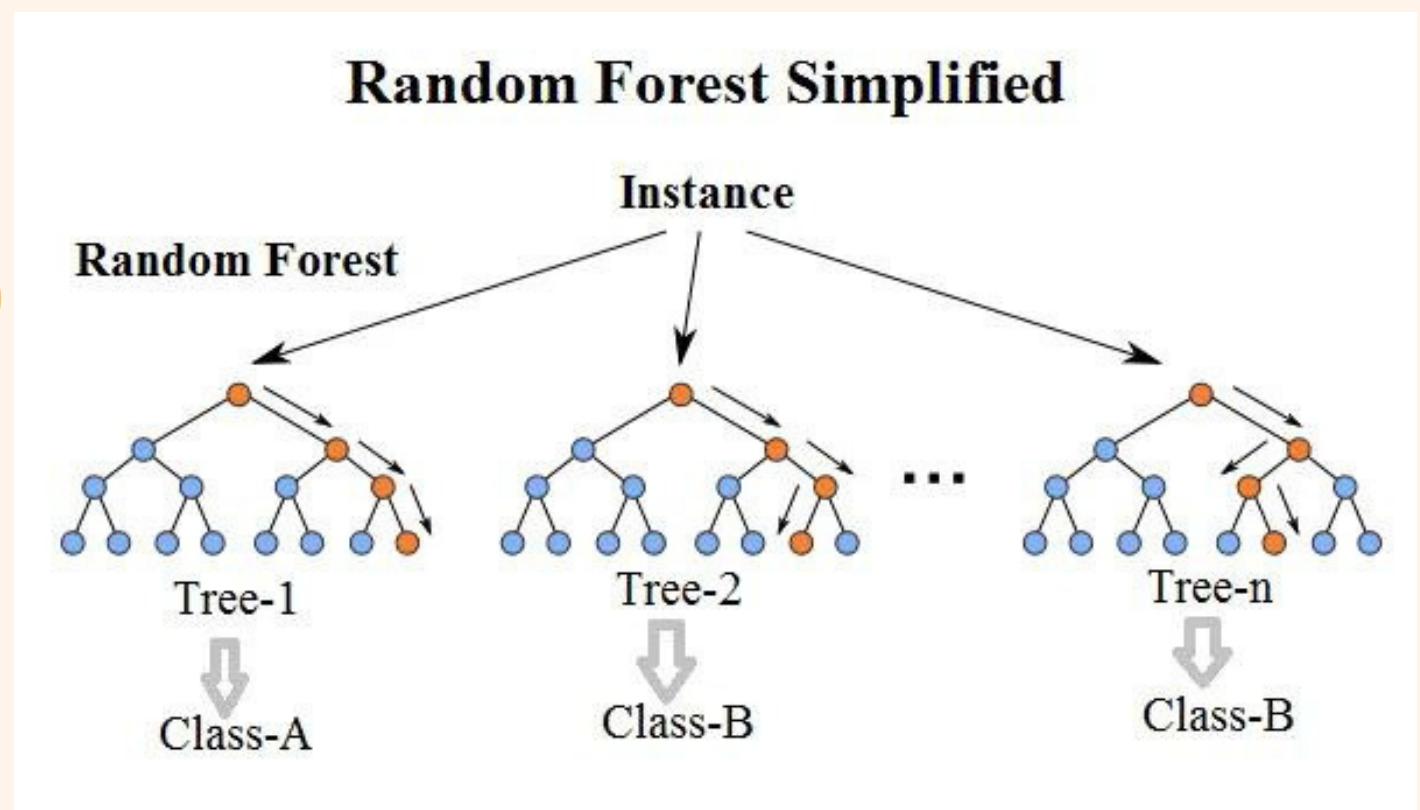
決策樹



<https://medium.com/jameslearningnote/%E8%B3%87%E6%96%99%E5%88%86%E6%9E%90-%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E7%AC%AC3-5%E8%AC%9B-%E6%B1%BA%E7%AD%96%E6%A8%B9-decision-tree-%E4%BB%A5%E5%8F%8A%E9%9A%A8%E6%A9%9F%E6%A3%AE%E6%9E%97-random-forest-%E4%BB%8B%E7%B4%B9-7079b0ddfbda>

Random forest

「三個臭皮匠勝過一個諸葛亮」

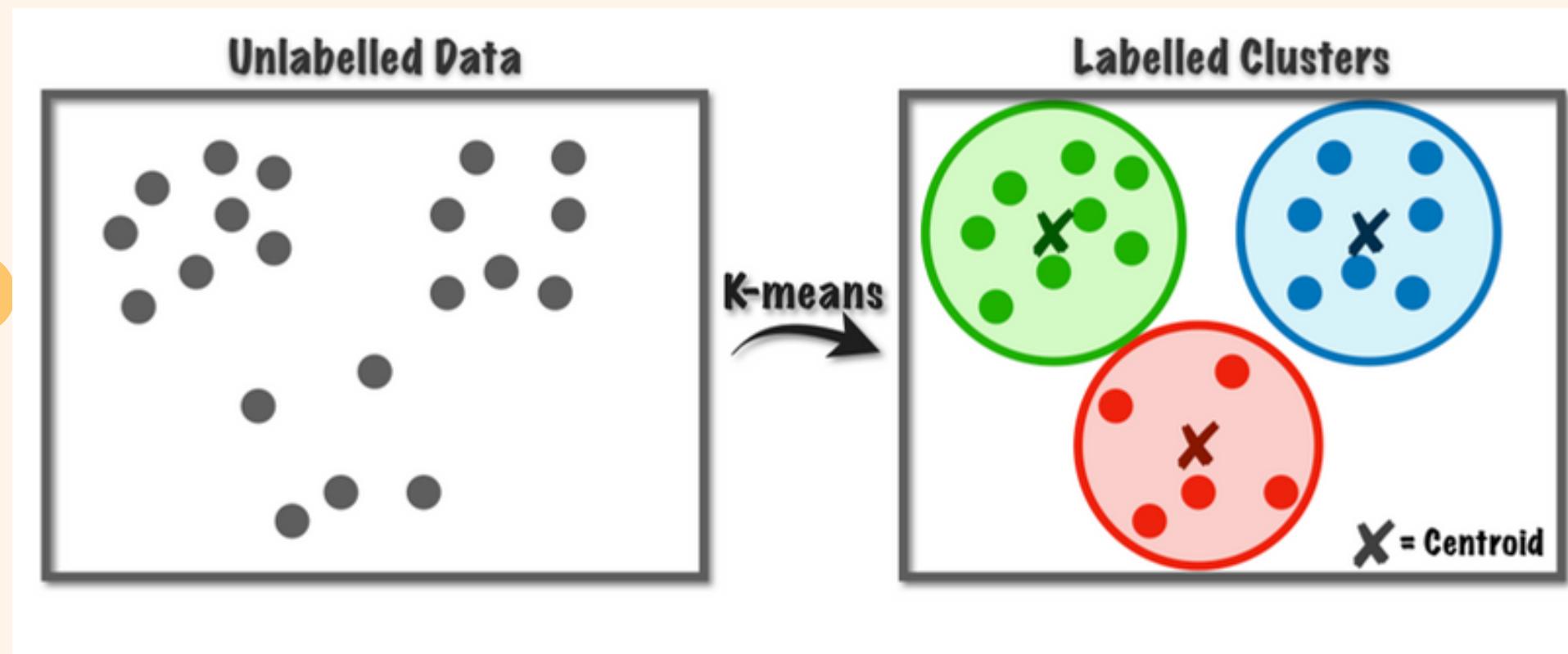


<https://medium.com/chung-yi/ml%E5%85%A5%E9%96%80-%E5%8D%81%E4%B8%83-%E9%9A%A8%E6%A9%9F%E6%A3%AE%E6%9E%97-random-forest-6afc24871857>

K-means

「物以類聚」

跟前面不一樣!?



https://pyecontech.com/2020/05/19/k-means_k-medoids/



交叉驗證 Cross-Validation

(補充 知道有這件事就可以)

比喻成kahoot

每堂課後做測驗看成效



超參數調整

hyperparameter tuning

(補充 知道有這件事就可以)

資料中: 特徵(欄位)&標籤(類別)->權重參數

模型本人: KNN的鄰居數->超參數

總結

今天的內容是機器學習中的一些算法
KNN，SVM，決策樹，隨機森林
K-means
除此之外還有很多很多
留給同學自行學習

觀念複習：
監督、非監督
分類、分群
訓練集、驗證集、測試集
Precision, recall, and accuracy

