

課程簡介

陳仁政

簡歷

- 中原大學電子工程所博士班
- 中央研究院博士後研究員
- 冠捷科技正工程師
- 104人力銀行資料科學家
- 鈔票辨識
- 電視韌體
- 評測資料分析與演算法
- 客群分布分析
- 各類影像辨識的應用系統
 - 交通執法、文件辨識等
 - 嵌入式與整合性系統開發

資料科學家的類別

- 開發新演算法
- 應用系統
 - 影像、聲音辨識
 - 自然語言分析
 - 工廠自動化
 - 自動化駕駛
 - 推薦系統
- 數據分析
 - 客群分布
 - 社群分析

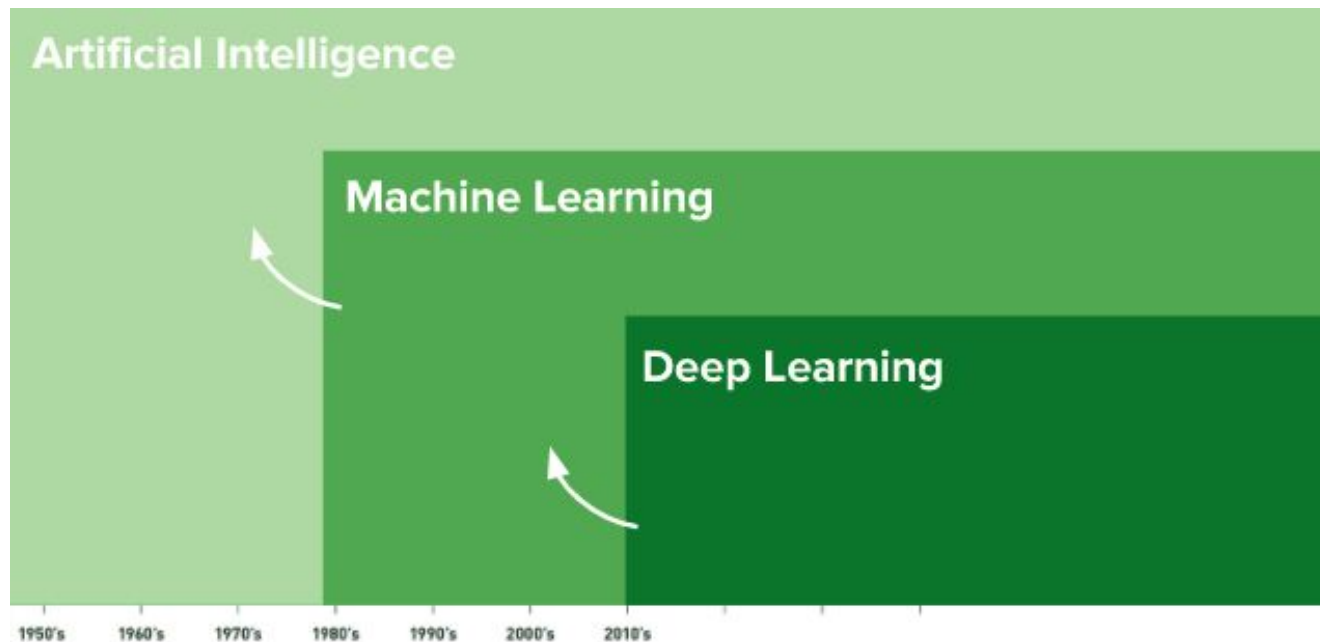
數據分析的步驟

- 資料收集
- 資料清洗
- 資料轉換
- 建立與選擇模型
- 評估結果
- 視覺化

資料科學需要的專長

- 程式設計
- 理論
 - 數理統計
 - 機械學習
- 領域知識

人工智慧、機器學習、深度學習？



什麼是機器學習？

- <https://zh.wikipedia.org/wiki/%E6%9C%BA%E5%99%A8%E5%AD%A6%E4%B9%A0>
- 機器學習是人工智慧的一個分支。
- 人工智慧的研究歷史有著一條從以「推理」為重點，到以「知識」為重點，再到以「學習」為重點的自然、清晰的脈絡。顯然，機器學習是實現人工智慧的一個途徑，即以機器學習為手段解決人工智慧中的問題。
- 機器學習在近30多年已發展為一門多領域交叉學科，涉及概率論、統計學、逼近論、凸分析、計算複雜性理論等多門學科。
- 機器學習理論主要是設計和分析一些讓電腦可以自動「學習」的演算法。機器學習演算法是一類從資料中自動分析獲得規律，並利用規律對未知資料進行預測的演算法。因為學習演算法中涉及了大量的統計學理論，機器學習與推斷統計學聯絡尤為密切，也被稱為統計學習理論。演算法設計方面，機器學習理論關注可以實現的，行之有效的學習演算法。很多推論問題屬於無程式可循難度，所以部分的機器學習研究是開發容易處理的近似演算法。

機械學習常見演算法

- 監督式

- 回歸

- linear regression
 - neural network
 - SVM

- 分類

- decision tree
 - k-NN
 - naive bayes
 - neural network
 - logistic Regression
 - SVM

- 非監督式

- 聚類

- k-mean
 - DBSCAN

Python 程式語言課程內容

- 語法
- 常用函式庫
- 資料庫
- 視窗程式設計
- 靜態網頁的網路爬蟲

人工智能實務課程內容

- 統計與機械學習基礎理論
- NumPy -- 數學陣列
- Pandas -- 數據
- 資料圖表視覺化
- scikit-learn -- 常用的機械學習函式
- Kaggle 案例分析

參考書籍

- **教課書**

- 精通 Python: 運用簡單的套件進行現代運算
 - <https://www.books.com.tw/products/0010858475>
- Python資料科學學習手冊
 - <https://www.books.com.tw/products/0010774364>

- **參考書籍**

- 精通機器學習: 使用Scikit-Learn, Keras與TensorFlow
 - <https://www.books.com.tw/products/0010854043>

聯絡方式

clement1972@gmail.com

課程資料

<https://bit.ly/36XVEDx>

使用程式

- Python
 - <https://www.python.org/downloads/>
- Visual Studio Code
 - <https://code.visualstudio.com/download>
- PSPad
 - <http://www.pspad.com/en/download.php>
- MySQL
 - <https://dev.mysql.com/downloads/mysql/>
- DB Browser for SQLite
 - <https://sqlitebrowser.org/dl/>

文件連結

- Python教學手冊
 - <https://docs.python.org/zh-tw/3/tutorial/index.html>
- Pandas
 - https://pandas.pydata.org/pandas-docs/stable/user_guide/index.html
- NumPy
 - <https://numpy.org/doc/stable/>
- HTML
 - <https://www.w3schools.com/html/default.asp>
- Beautiful Soup
 - <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- SQL
 - <https://www.w3schools.com/sql/default.asp>

Kaggle

- <https://www.kaggle.com/>
- <https://www.kaggle.com/uciml/iris>
- <https://www.kaggle.com/c/titanic>

Google Colab

<https://colab.research.google.com/>

書本範例

- <https://github.com/ageron/handson-ml2>
- <https://github.com/jakevdp/PythonDataScienceHandbook>
- <https://github.com/madscheme/introducing-python>