

碩二上 上課時間：星期二下午

# 機器學習

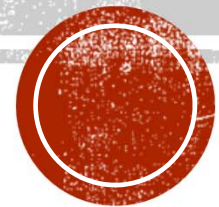
# MACHINE LEARNING

課程介紹 吳智鴻教授

國立臺中教育大學 數位內容科技研究所

<http://120.108.221.55/profchwu/ml>

Email: [chwu@mail.ntcu.edu.tw](mailto:chwu@mail.ntcu.edu.tw)

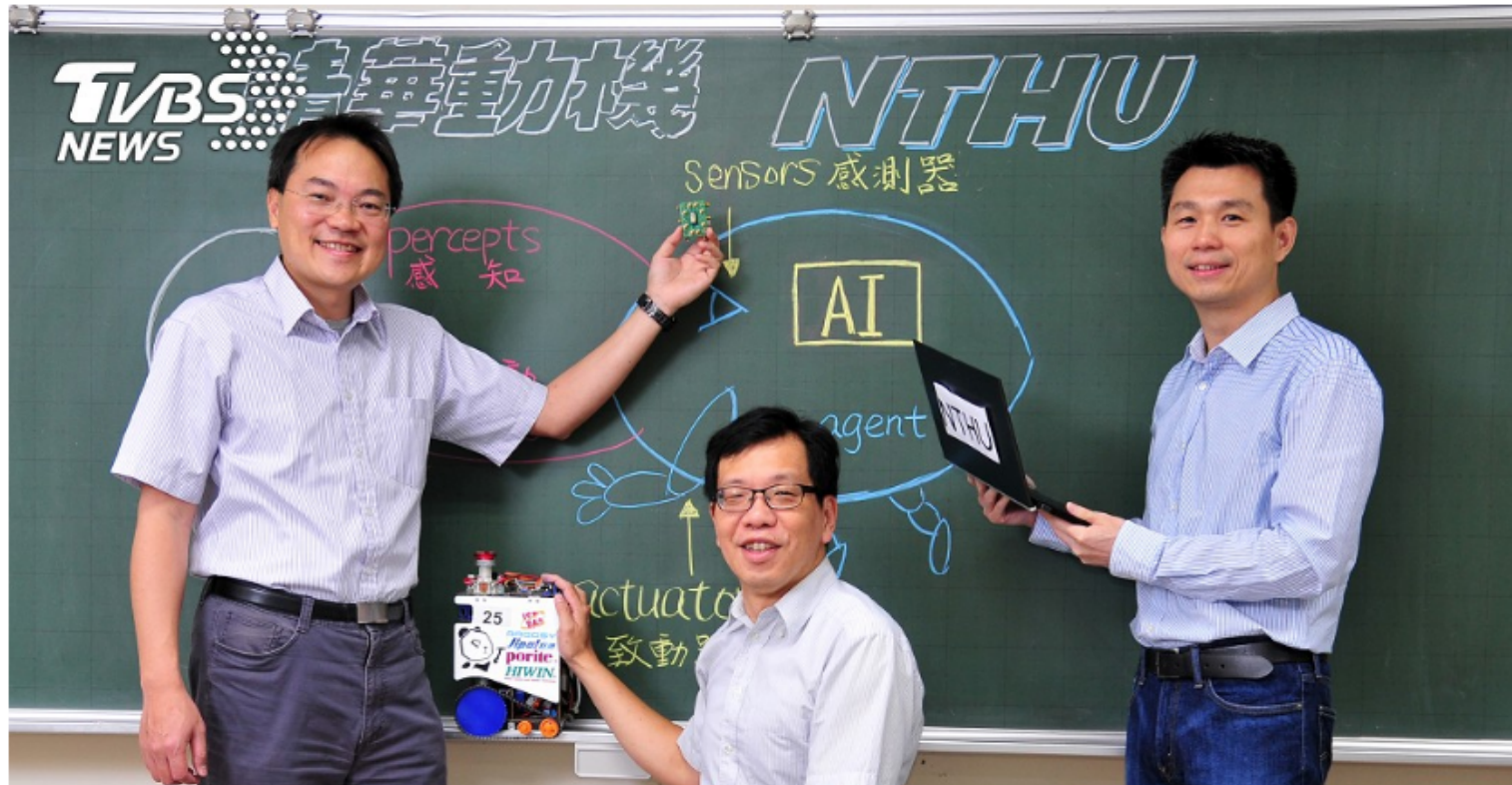


# 課程的重要性

- 大數據的時代，我們面臨許多資料，但卻不知道如何分析？
- 學會這個技術，有利於後續進入AI的領域。
- 目前各行各業都很亟需大數據的分析人才。



# 清華動機整合光機電 再添 AI 新能量！欲培訓下一代所需科技人才



▲清華大學欲培訓下一代所需科技人才。(圖／翻攝自中央社)

2018-04-02 15:39:41

讚 98 分享 G+

清華動機系今年2月禮聘兼具機械及資訊跨領域專長的丁川康教授，開設「人工智慧」課程。蔡宏營表示，動機系核心課程將結合人工智慧與資訊科學，包括**最佳化技術**、**機器學習**、**資料探勘與大數據分析**，推動機械工業與製造生產系統的智慧化，培育「實用的AI人才」。



# 國內外長春藤名校均開設此課程

## 机器学习

**About this course:** 机器学习是一门研究在非特定编程条件下让计算机采取行动的学科。最近二十年，机器学习为我们带来了自动驾驶汽车、实用的语音识别、高效的网络搜索，让我们对人类基因的解读能力大大提高。当今机器学习技术已经非常普遍，您可能在毫无察觉情况下每天使用几十次。许多研究者还认为机器学习是人工智能（AI）取得进展的最有效途径。在本课程中，您将学习最高效的机器学习技术，了解如何使用这些技术，并自己动手。

More

Created by: 斯坦福大学

Stanford University



**Taught by:** Andrew Ng, Co-founder, Coursera; Adjunct Professor, Stanford University; formerly head of Baidu AI Group/Google Brain

Language	English, Subtitles: Spanish, Hindi, Japanese, Chinese (Simplified)
How To Pass	Pass all graded assignments to complete the course.
User Ratings	★★★★☆ Average User Rating 4.9 <a href="#">See what learners said</a>

## 机器学习基础：案例研究

**About this course:** 你是否好奇数据可以告诉你什么？你是否想在关于机器学习促进商业的核心理解？你是否想同专家们讨论关于回归，分类，深度学习以及推荐系统的一切？在这门课上列实际案例学习来获取实践经验，在这门课程结束的时候，

Created by: 华盛顿大学



**Taught by:** Carlos Guestrin, Amazon Professor of Machine Learning Computer Science and Engineering



**Taught by:** Emily Fox, Amazon Professor of Machine Learning Statistics

## 机器学习 A-Z (Machine Learning A-Z in Chinese)

全面建立机器学习的知识架构，并且在Python和R里构建不同的机器学习模型。课程内容包括所有的代码模板。

畅销课程 ★★★★★ 4.7 (315 个评价) 1938 名學生註冊

建立者: Hadelin de Ponteves, Kirill Eremenko, 武亦文 Yiwen, 李秦 Qin 上次更新 4/2018

Simplified Chinese

### 我會學些什麼呢？

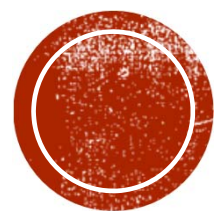
- 完全掌握机器学习及在Python和R里的应用
  - 做出准确的预测和强大的分析
  - 利用机器学习解决私人问题
  - 掌握并熟练处理先进的技术，例如对降低数据维度
  - 建立起强大的机器学习知识架构，并且知道如何创建和运用不同的模型来解决任何问题
  - Have a great intuition of many Machine Learning models
- 深刻理解各种机器学习的模型
  - 利用机器学习创造更多价值
  - 掌握并熟练处理强大的算法，例如强化学习，自然语言处理，还有深度学习
  - 了解对不同的问题怎样选择合适的机器学习模型
  - Master Machine Learning on Python & R
  - Make accurate predictions and powerful

檢視更多



# 機器學習就像是武俠小說裡深厚的內功





# 為什麼要學機器學習？



是一門資料分析的  
技術



把資料萃取出有價值的資訊  
把資訊加值





# 課程會學到什麼

## 學習內容

- 兩大機器學習語言 R, Python
- Weka 視覺化分析軟體
- 大數據資料處理技術
- 資料分析
  - 統計量分析、
  - 購物籃分析、
  - 異常值偵測
  - 迴歸分析
  - 決策樹分析、類神經網路分析等 )、
- 基礎機器學習建模與評估。
- 並導覽進階的互動式視覺化、深度學習、自然語言處理等 Python 技術套件。

## 應用案例

- 文件分類
- 金融資料預測
- 情緒分析
- ptt 網路爬蟲技術
- 文字雲技術



# 上課內容與方式

## 分析工具

- 不需要程式很強，也可以學習
- R與Python程式的說明與提供
- 視覺化WEKA分析軟體的介紹
- 實際 上機操作與分析

## 分析內容

- 分析方法的講解
- 分析技術與原理的講解
- 以實際案例資料集進行分析
- 實際 資料集的提供



# WHAT IS MACHINE LEARNING?

- Learning from Data
- 是一門教導你如何有效分析你的資料的學科
- 也是AI人工智慧的前導課程



# 機器學習與學習的區別

## ■ 人類



## ■ Machine Learning



# 機器學習

幫助你對資料作更好的分析，提高分析的準確度

- 分類
  - 股票資訊 -> 機器學習ML -> 股市漲跌
- 預測
  - 股票資訊 -> 機器學習ML -> 股票價格預測
- 辨識
  - 車牌辨識、人臉辨識等
- 分群
- 關連      購物籃分析



# MACHINE LEARNING 可以應用的產業

- 教育/數位學習
- 商業
- 醫學
- 工程
- 科學
- 化學
- 遊戲
- .....



# 與統計學的不同

- 統計學是分析資料，並對資料提出假設與驗證。
- 機器學習方法的優點：
  - 不需要對資料型態作一些假設（更符合真實世界的狀況）
  - 不需要提出假設
  - 少量的資料就可以進行分析（不需要大樣本）
  - 通常預測出來的準確度較高
  - 有許多功能是統計學中沒有的分析



# 機器學習可協助人工智慧

Learning from Data      What is Machine Learning

## Machine Learning and Artificial Intelligence

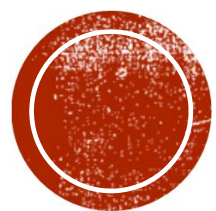
<b>Machine Learning</b>	<b>Artificial Intelligence</b>
use data to compute something that improves performance	compute <b>something that shows intelligent behavior</b>

- **improving performance** is something that shows **intelligent behavior**
  - **ML can realize AI**, among other routes
- e.g. chess playing
  - traditional AI: game tree
  - ML for AI: 'learning from board data'

ML is one possible **and arguably most popular** route to realize AI







# 產業的應用



# 你知道嗎？

- 電腦如何戰勝世界棋王？Google如何運用關鍵字預測流感爆發？
- Twitter如何從發文中預測Facebook的股價？
- 銀行如何即時防範信用卡盜刷？  
這些靠的全是Big Data巨量資料系統分析！



# 機器學習與教育雲

- 透過學研界在教育雲大數據之研究創意與能量
  - 萃取教育雲的可用資訊
  - 找出教育雲的資料價值
  - 找出教育雲的應用情境
  - 善用大數據分析的結果
  - 協助解決社會與產業關注的問題



# 一些已經應用的商業領域 車牌辨識

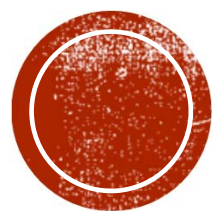


# 語音辨識



# 臉部辨識





# 職場需求大



# 21世紀資訊業最酷炫、最誘人的職業——資料科學家！

- 巨量資料分析與企業智慧系統服務的人才需求，在2015年全球將創440萬份工作。  
根據《[財訊雜誌 第419期](#)》的文章報導，「資料科學家入門就領44K」  
要成為資料科學家，需具備軟體開發技術、資料處理分析和專業領域知識等技能，  
這些全部都會教給你！讓你步入資料科學領域的大門，跟上時代的潮流！



讓資料會說話，就是「資料科學家」的特色，也是巨量資料**BIG DATA**的精髓。

- 國際研究暨顧問機構**Gartner**表示：2013年，64%企業已經或計劃投資巨量資料！
- 根據《**Information Week**》所做的「巨量資料人才爭奪戰」趨勢報告預估到2018年資料分析專業人才將爆增170萬名的專業資料分析人才缺口！
- 81%台灣企業認同巨量資料有助做出更好決策，並且相信未來三年，資料中心自動化將是創新與快速反應的關鍵。





# 市場職缺 需求大

<input type="checkbox"/>	01/04	<b>數據分析師</b>	福建省廈門市	<a href="#">我要應徵</a>
		 <a href="#">玉晶光電股份有限公司</a>		<a href="#">6-10人應徵</a>
		工作經歷不拘，碩士學歷，待遇面議 1.企業資訊系統 <b>數據分析</b> 2 指導使用者單位使用大 <b>數據技術分析</b> 問題		
<input type="checkbox"/>	03/16	<b>數據分析師</b>	台北市內湖區	<a href="#">我要應徵</a>
		 <a href="#">創星物聯科技股份有限公司</a>		<a href="#">11-30人應徵</a>
		2年以上工作經歷，大學學歷，待遇面議 (1). <b>數據整理及分析</b> 。使用機器學習演算法，統計理論及方法 <b>分析</b> 物聯網、車聯網使用者行為資料。(2). <b>分析和詮釋數據</b> ...		
<input type="checkbox"/>	03/08	<b>07_數據分析專員</b>	新北市中和區	<a href="#">我要應徵</a>
		 <a href="#">台灣之星電信股份有限公司</a>		<a href="#">11-30人應徵</a>
		工作經歷不拘，大學學歷，待遇面議 1. 解析電信資料庫 <b>數據</b> 及使用者行為 <b>分析</b> 。 2. 從 <b>數據庫</b> 發掘使用者行為模式與偏好，提出分群描述及預測模型。 3. <b>分析報告製</b> ...		
<input type="checkbox"/>	03/15	<b>數據分析師</b>	台中市西屯區	<a href="#">我要應徵</a>
		 <a href="#">普拉瑞斯創意整合有限公司</a>		<a href="#">0-5人應徵</a>
		3年以上工作經歷，高中學歷，待遇面議 1.資料庫 <b>數據</b> 及使用者行為 <b>分析</b> 2.競爭對手 <b>分析</b> ，協助產品做針對性的調整 3.負責整合、規範和整理產品 <b>數據庫</b> ，為建立[[ <b>數據</b> ]]...		
<input type="checkbox"/>	03/17	<b>數據/市場分析師</b>	台北市松山區	<a href="#">我要應徵</a>
		<a href="#">伊凡達科技股份有限公司</a>		<a href="#">11-30人應徵</a>
		2年以上工作經歷，專科學歷，待遇面議 1.遊戲資料庫 <b>數據</b> 及使用者行為 <b>分析</b> 。 2. <b>數據分析</b> 使用者行為模式與偏好，會員分群經營與消費行為 <b>分析</b> 。 3.會員與品牌定...		
<input type="checkbox"/>	03/09	<b>數據/市場分析師</b>	台中市西屯區	<a href="#">我要應徵</a>
		<a href="#">靈識有限公司</a>		<a href="#">0-5人應徵</a>
		2年以上工作經歷，專科學歷，待遇面議 1.遊戲資料庫 <b>數據</b> 及使用者行為 <b>分析</b> 。 2. <b>數據分析</b> 使用者行為模式與偏好，會員分群經營與消費行為 <b>分析</b> 。 3.會員與品牌定...		

# 高額獎金&廠商人才需求

Home Competitions Discussion Datasets [Sign In](#)

## 台灣ETF價格預測競賽

進行中

[Overview](#) [Leaderboard](#) [Download Dataset](#)

### 比賽獎金

新台幣 12 萬元，並提供趨勢科技預聘書	第一名
新台幣 5 萬元	第二名
新台幣 3 萬元	第三名

### 競賽說明

股價是否能被預測，一直存在著正反兩派的觀點。有效市場派認為股價波動是隨機遊走而無法預測下一步會怎麼走，往那個方向走。然而有效市假說似乎又與現實相悖，許多研究顯示短期股價的相關性並不為零，會有動量的存在，而這是趨勢投資的基石。

因此我們提供參賽者 台灣上市/櫃公司近五年來的歷史每日股價(開盤，最高，最低,收盤)及成交量，希望能集合各位參賽者的智慧來找出預測股價的模型與方法。

參賽者依主辦單位所提供台灣上市/櫃近五年來的歷史每日股價(名目股價與調整股價)及成交量來預測台灣十八檔上市櫃成分證券ETF在下一週五天的漲跌及價格。



79  
參賽隊伍



總獎金  
新台幣 20 萬元

開始 4/3/2018 結束 6/22/2018



# 產業需求大

- 金融業
- 醫療業
- 教育業
- 工業
- .....

全部工作、關鍵字(資料科學)

相關搜尋： 巨量資料、人工智慧、語意分析、自然語言處理、資料探勘

修改條件

## 3/22 資料科學工程師

昱冠資訊股份有限公司 | 電腦軟體服務業

新竹市 | 經歷不拘 | 大學

昱冠資訊使用資料科學方法(統計建模 Statistical Model/機器學習 Machine Learning/深

面議 無經驗可

★ 儲存

✉ 應徵

11~30人應徵

## 4/03 RS180222-資料科學應用顧問

企 緯創資通股份有限公司 | 電腦及其週邊設備製造業

新北市汐止區 | 5年以上 | 大學

1. 與客戶溝通討論，協助客戶資料的萃取及整理 2. 協助參與客戶資料的分析與問題，假

面議 上市上櫃 員工90000人

★ 儲存

✉ 應徵

0~5人應徵

## 4/03 RS180222-工業 4.0 資料科學顧問

企 緯創資通股份有限公司 | 電腦及其週邊設備製造業

新北市汐止區 | 經歷不拘 | 大學

1. 與客戶溝通討論，協助客戶資料的萃取及整理 2. 協助參與客戶資料的分析與問題，假

面議 無經驗可 上市上櫃 員工90000人

★ 儲存

✉ 應徵

0~5人應徵

## 4/03 RS180313-醫療資料科學家

企 緯創資通股份有限公司 | 電腦及其週邊設備製造業

新北市汐止區 | 5年以上 | 碩士

1. 能獨立運作，與客戶溝通討論，協助客戶釐清需求，提出資料運用模型 2. 透過統計，

面議 上市上櫃 員工90000人

★ 儲存

✉ 應徵

0~5人應徵

## 4/09 資料科學顧問

企 勤業眾信聯合會計師事務所 | 會計服務業

台北市中山區 | 經歷不拘 | 大學

1. 提供企業各類資料分析需求解決方案，包含分析方法規劃、建模、預測分析、視覺化

面議 無經驗可 員工3600人

★ 儲存

✉ 應徵

6~10人應徵

## 4/03 資料科學工程師

企 玉山銀行\_玉山商業銀行股份有限公司 | 銀行業

台北市中山區 | 經歷不拘 | 大學

★ 儲存



# 電信業

## 07\_數據分析專員

台灣之星電信股份有限公司 本公司其他工作

1. 解析電信資料庫數據及使用者行為分析。
2. 從數據庫發掘使用者行為模式與偏好，提出分群描述及預測模型。
3. 分析報告製作並提出觀點與提案。
4. 其他主管交辦事項

職務類別：[其他專案管理師](#)、[市場調查／市場分析](#)、[行政人員](#)

工作待遇：面議

工作性質：全職

上班地點：[新北市中和區](#)

管理責任：不需負擔管理責任

出差外派：無需出差外派

上班時段：日班，09:00-18:00

休假制度：週休二日

可上班日：一個月內

需求人數：1 人

### 條件要求

接受身份：上班族

工作經歷：不拘

學歷要求：大學、碩士

科系要求：不拘

語文條件：[英文](#) -- 聽 / 略懂、說 / 略懂、讀 / 略懂、寫 / 略懂

擅長工具：不拘

工作技能：不拘

其他條件：具備1年以上數據分析相關實務經驗

# 遊戲產業

## 數據分析師

傳奇網路遊戲股份有限公司 本公司其他工作

### 工作內容

手機遊戲營運數據分析  
建置遊戲內經濟活動相關模型

職務類別：[其他資訊專業人員](#)

工作待遇：依公司規定

工作性質：全職

上班地點：[台北市信義區市民大道6段288號7樓](#)

管理責任：不需負擔管理責任

出差外派：無需出差外派

上班時段：日班，09:30~18:30

休假制度：週休二日

可上班日：不限

需求人數：2 人

### 條件要求

接受身份：上班族、應屆畢業生

工作經歷：1年以上

學歷要求：大學、碩士

科系要求：不拘

語文條件：不拘

擅長工具：不拘

工作技能：不拘

其他條件：對數字敏感度及邏輯分析能力

熟悉排列組合、機率理論

具企業經營分析或遊戲企劃經驗



# 行銷業

人力銀行 My104 家教 外包 獵才 教育 職涯社群 更多 ▾ 「超級業務講堂」免費報名中

## 行銷數據分析師

台灣寶生堂股份有限公司 本公司其他工作


### 工作內容

1. 資料庫數據及使用者行為分析
2. 數據分析使用者行為模式與偏好和消費者行為分析
3. 會員及品牌定位解析及關連度解析
4. 透過顧客資料解析,提出對公司在開發顧客及客戶經營上的規劃及建議
5. 掌握數位平台發展與創新應用趨勢,提出對公司有利的創新解決方案及專案規劃

職務類別: [市場調查](#)/[市場分析](#)

工作待遇: 面議

工作性質: 全職

上班地點: 台北市大安區敦化南路一段111號11樓 

管理責任: 不需負擔管理責任

出差外派: 需出差,一年累積時間未定

上班時段: 日班

休假制度: 依公司規定

可上班日: 可年後上班

需求人數: 1 人

### 條件要求

接受身份: 上班族

工作經歷: 3年以上

學歷要求: 大學、碩士

科系要求: 不拘

語文條件: 英文 -- 聽/精通、說/精通、讀/精通、寫/精通

日文 -- 聽/中等、說/中等、讀/中等、寫/中等

擅長工具: 不拘

# 數位創意產業

人力銀行 My104 家教 外包 獵才 教育 職涯社群 更多 ▾ 「超級業務講堂」免費報名中

## 04 人力銀行

## 數據分析師

宏麗數位創意股份有限公司 本公司其他工作

### 工作內容

1. 資料庫數據及使用者行為分析。
2. 數據分析使用者行為模式與偏好和消費行為分析、提出對客戶體驗相關優化改善措施。
3. 會員與品牌定位解析及關連度解析。
4. 針對異常數據提出檢討及稽核。
5. 蒐集外部競爭情勢之現況調查、彙整分析及歸納建議。



職務類別: [市場調查](#)/[市場分析](#)

工作待遇: 面議

工作性質: 全職

上班地點: 台北市中山區南京東路三段70號8樓 

管理責任: 不需負擔管理責任

出差外派: 無需出差外派

上班時段: 日班, 09:00~18:15

休假制度: 週休二日

可上班日: 一個月內

需求人數: 1 人

### 條件要求

接受身份: 上班族

工作經歷: 2年以上

學歷要求: 專科、大學、碩士

科系要求: 不拘



# 工研院

## 資科中心\_資訊系統開發與數據分析工程師(F200)

工研院\_財團法人工業技術研究院 本公司其他工作

負責資訊系統開發(系統分析、設計及程式撰寫)及數據分析應用。

職務類別：[軟體設計工程師](#)、[電腦系統分析師](#)、[Internet程式設計師](#)

工作待遇：面議

工作性質：全職

上班地點：[新竹縣竹東鎮中興路四段195號](#)

管理責任：不需負擔管理責任

出差外派：需出差，一年累積時間未定

上班時段：日班

休假制度：依公司規定

可上班日：一個月內

需求人數：1 人

### 條件要求

接受身份：上班族、應屆畢業生

工作經歷：不拘

學歷要求：大學、碩士

科系要求：數學及電算機科學學科類

語文條件：不拘

擅長工具：不拘

工作技能：不拘

其他條件：1.資科、資工、資管等相關科系碩士以上。

2.具資訊系統開發、程式撰寫相關經驗。

3.熟悉數據分析(Data Analytics)、機器學習(Machine Learning)。

4.熟悉MS SQL、C#、APP、R/Python程式語言。

# 遊戲產業

104人力銀行

## 遊戲市場數據分析師

[創越數位行銷有限公司](#) 本公司其他工作

### 工作內容

- 1.解析遊戲平台資料庫數據及使用行為分析。
- 2.從數據庫發掘使用者行為模式與偏好，提出分群描述及預測模型。
- 3.消費者與品牌定位解析及關連度解析。
- 4.資料視覺化呈現。
- 5.提出觀點與提案，為各部門提供資料解析支援及決策應用依據。
- 6.針對異常數據提出檢討及稽核
- 7.蒐集外部競爭情勢之現況調查、彙總分析及歸納建議

另需要具備以下工作能力：

- 1.對網路產業有極高度興趣，曾有數據分析/挖掘、數位媒體行銷、或產品行銷相關工作經驗者尤佳。
- 2.對資料分析有高度敏感性與熱情，熟悉數據處理相關技術或工具者，如：SPSS, R 語言，並能提案觀點加以應用者。

職務類別：[市場調查](#)/[市場分析](#)

工作待遇：面議

工作性質：全職

上班地點：[台中市西屯區市政北二路236號15樓之1](#)

管理責任：不需負擔管理責任

出差外派：無需出差外派

上班時段：日班

休假制度：週休二日

可上班日：不限

需求人數：1 人



# 評分方式

- 平時成績 30%
- 期中分析作業 30%
- 期末專題報告 40%
  - 以實際的資料集，進行分析報告



# 延伸課外閱讀 ~ 不錯的教材

53 people clipped this slide

Quick Tour of Machine Learning  
(機器學習速遊)

Hsuan-Tien Lin (林軒田)

Chief Data Scientist @ Appier (沛星互動科技)

**Appier**

資料科學愛好者年會系列活動, 2017/03/11

Hsuan-Tien Lin (Appier)

Quick Tour of Machine Learning

0/128

1 of 164





# 延伸課外閱讀 ~ 不錯的教材

37 people clipped this slide

## 推薦系統的介紹

林守德教授

CSIE, NTU

[sdlin@csie.ntu.edu.tw](mailto:sdlin@csie.ntu.edu.tw)

1

◀ 1 of 219 ▶



# 延伸課外閱讀 ~ 不錯的教材

70 people clipped this slide

資料科學系列活動一

使用 R 語言建立自己的**演算法交易事業**

Speaker: 吳牧恩

Date: 2016.03.11(六)

2017/03

1

◀ 1 of 211 ▶

