

# 淺談深度學習攜手大 數據引爆的第四次工 業革命—AI 人工智慧

劉毓容

## 前言

2016 年 Google 開發的人工智慧（以下簡稱 AI）程式 AlphaGo 擊敗世界圍棋冠軍、2017 年 Open AI 公司的 AI 程式也在世界級的電玩遊戲競賽中打敗人類對手，凸顯 AI 發展潛力驚人，未來不僅將成為全球最大商機之一，日常生活的應用範圍也將不斷擴大，如銀行的智能服務、AI 叫車平台、聊天機器人等，昔日科幻的電影情節，如今已真實上演，顯見 AI 時代的來臨！AI 不但是科學技術的革新，也將全面性改變人類的生活與產業。舉凡零售、金融服務、生技醫療、自動駕駛、機器人、智慧生活等，都是 AI 應用的領域，衍生的商機無窮，亦引發全球發展 AI 之風潮。政府為掌握 AI 發展的契機，繼宣示 106 年為台灣 AI 元年後，續於同年 8 月推出「AI 科研戰略」，並於 107 年 1 月 18 日起推動 4 年期的「台灣 AI 行動計畫」（107 年至 110 年），全面啓動產業 AI 化，有鑑於此，本文由深度學習與大數據等組成元素，探討 AI 人工智慧相關的多元應用與發展。

## 壹、AI 人工智慧（Artificial Intelligence, AI）與電子商務

人工智慧（Artificial Intelligence, AI）歷經多年的發展，已形成一股不容忽視的技術浪潮，Google 於 2015 年底公布了 TensorFlow 開源軟體庫，至 2016 年已成為受歡迎的機器學習開園專案，支援了各種深度學習（Deep Minging）的演算法，尤其 2016 年 Google DeepMind 開發了人工智慧圍棋程式 AlphaGo，並成功於 2016 年戰勝棋王，顯示了人工智慧在不久的未來凌駕於人類之可能性。以下就人工智慧市場概況與產業運用做說明。

### 一、人工智慧市場概況

人工智慧每天皆非常迅速成長，除了簡化醫療保健行業的工作流程、減少了教育部門行政任務的工時外，同時也降低了製造業的管理成本，由 Adobe 研究中，現在約有 15% 企業正在使用人工智慧，也預計未來 1 年（2020 年）將會成長 31% 左右，更統計了 2013 年以後至今，需要人工智慧的工作需求增加了約 450%，由於網路的普及，大數據及機器學習的人工智慧持續的發展與進步，使人工智慧在近幾年有更大的技術突破，其中，人工智慧技術包含人工神經網絡（Deep Learning）、大數據的機器學習（Machine Learning）以及計算

機視覺和語言，而計算機視覺和語言包含自然語言處理（Nature Language Processing, NLP）、語音辨識與語意技術，由圖 1-1 可知，機器學習（Machine Learning）吸引約 50 億至 70 億元美金的投資金額，約佔了投資金額的 60%，投資意願相當高。金融產業不僅結合創新科技，且創造新商業模式，使金融創新技術越來越受矚目，其中，金融創新技術主要針對金融機構中勞動密集性、重複性較高，或需要處理大量資料的繁雜作業，較無效率的工作流程進行「去中介化」，有助於減少傳統人力資源，並透過機器縮短加工生產的週期，並取代人力費時的操作過程，除此之外，金融創新技術亦可將作業風險降低，改變了傳統金融業者的獲利結構，改變全球的零售金融業，包含第三方支付、小額借貸、投資與財富管理等。

2018 年 8 月 Juniper Research 預測，零售業在 2022 年投資於機器學習的成本會大幅提升將近四倍，若零售業也致力投資於人工智慧，能更改進對顧客的服務品質，包含提供顧客量身訂做的自動化行銷平台，不但可即時報價，亦能提供即時服務聊天之機器人，且研究將人工智慧分為三大類型，第一為弱人工智慧（Artificial Narrow Intelligence, ANI），其解決問題注重於較狹隘的領域與主題，再者為強人工智慧（Artificial General Intelligence, AGI），具

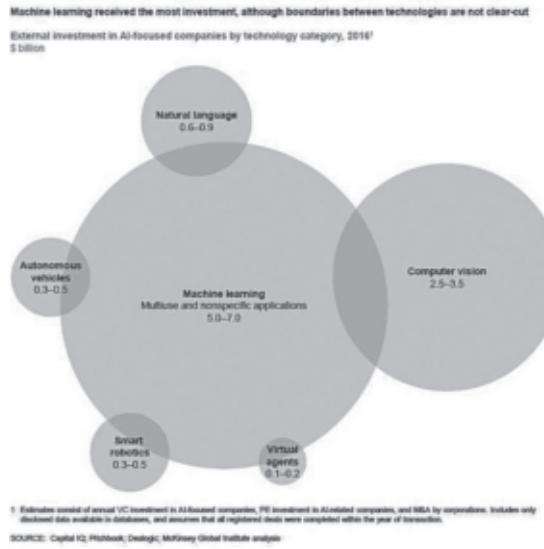


圖 1-1 2016 年人工智慧領域的外部投資  
(US\$ billion)



資料來源：McKinsey Global Institute  
(Bughin et al., 2017)

有廣泛解決問題的能力，其表現具有人類水準，最後為超級人工智慧（Artificial Super Intelligence, ASI），其表現不但具有人類水準，更可能在相同目標下贏過人類的表現。而零售業主要應用人工智慧的兩大關鍵趨勢，其一為個人化需求，在競爭激烈的環境中，若能提供一個較同業具有引人注目的差異化因素，與傳統的顧客服務相比，不僅能維持服務準確性，也能提供更好的服務品質，例如聊天機器人與目標行銷，更能為零售業達到策略目標；第二

關鍵則為奢侈品企業，由於小型零售業難以負擔人工智慧的高成本，使奢侈品企業為較早應用人工智慧之零售業，主因為奢侈品非常仰賴顧客的購買，其消費型態較不同於一般產品，所期望的服務亦不同，而運用人工智慧技術可提升顧客購買的機會，因此若能使人工智慧的技術更加普遍且成本較為低廉時，其他小型零售業也將能夠跟進運用人工智慧的技術。

## 二、電子商務零售業市場現況

因應全球化的趨勢、科技快速的發展使得電商通路的發展一日千里，造成傳統零售業面臨一大挑戰，在線零售銷售的顯著增長顯示了消費者購買和使用金融服務的方式已經發生了根本性的變化，與傳統零售商店中的購物方式相比，網上購物具有一些獨特的特點，例如在線零售之顧客的購物每個顧客的購物流程和活動都可以即時，準確地跟蹤，每個顧客的訂單通常與送貨地址和帳單地址相關聯，每個顧客有一個在線商店帳戶，提供必要的聯繫和付款信息。而透過在線購物的特徵使得在線零售商可以透過每個客戶之購買資訊，對顧客進行個性化的分析，並以客戶為中心的商業智能分析以提供行銷策略，在線零售業者時常關注的議題為客戶對哪些商品具有興趣，因此對於客戶瀏覽的產品網頁可提供零售業者相當重要的資訊，且客

戶在每個網頁上所停留的時間，抑或是瀏覽過哪一系列之商品網頁等資訊，皆可帶來消費者的資訊，並可從中挖掘出具有價值之顧客，以及其獨特之處，隨著顧客消費型態的改變，重心漸漸轉移至電子商務通路，使得零售業者不得不轉型以維繫故有之顧客，此外，新一代之消費者對於產品品牌之忠誠度相較於過去消費者低，其特色以及購買行為模式，更是在線零售業者欲關注之議題之一。在行銷分析中，亦可透過其消費行為以觀察產品之間之相依需求，意即有哪些產品顧客會同時購買，其購買產品之順序等，各種消費模式皆可謂在線零售業者帶來許多資訊。以零售業為例，中國阿里巴巴買下中國大潤發，關鍵因素為阿里巴巴旗下的電商平台天貓於2018年完成了「天貓新零售」的改造，其三大核心策略分別為門市數據化改造、多業態多通路發展、重新定義大賣場，天貓加上大潤發，線上與線下銷售接軌，大潤發聯手天貓、盒馬先生推出「1小時送達服務」以吸引門市所在地相近的年輕人；美國電商龍頭亞馬遜（Amazon）於2017年買下以高價有機食品聞名的美國全食超市（Whole Foods），走向新零售的趨勢，對於亞馬遜而言，過去即有推出Amazon Fresh的生鮮商品運送服務，藉由Whole Foods既有的商品網路進一步擴展旗下生鮮商品事業，另一方面，Whole Foods亦可透過亞馬

遜的資源與品牌聲譽，幫助企業與其零售商競爭者在此激烈的零售市場中脫穎而出。利用人工智慧以及大數據分析，讓製造業可更具服務化、客製化，更可預測消費者需求並提供量身訂做的商品，進行精準行銷與服務，使線上、線下以及物流整合為一體。

電商通路可藉由網路與顧客接觸，加上藉由人工智慧在大量數據中尋找特定的消費模式，讓業者可充分利用已知的顧客喜好特性，為未知顧客的偏好提供建議，透過人工智慧以及機器學習能夠利用銷售、消費者以及各種數據資料，以提升網路搜尋的功能，更可透過客戶資料以及顧客所接觸之商品屬性進一步預測消費者對新商品以及新風格的需求，美國電商龍頭亞馬遜（Amazon）已於2013年獲得「預期出貨（anticipatory shipping）」的專利，不僅考量了季節性的需求，亦考量天氣、人口統計以及獨特的顧客消費模式所帶來的影響，隨著人工智慧的快速發展，各行業領域透過大數據以及人工智慧技術的多元性來了解現代社會的需求，並藉由機器深度學習給予產業進行更精準的市場預測。

### 三、AI 人工智慧（Artificial Intelligence, AI）

#### （一）人工智慧（AI）的演進

自1958年美國賓州匹茲堡大學的兩位

人工智慧發展先驅 Herbert Simon 以及 Allen Newell 發表出已有機器可以自行思考、發展與創造，且機器之能力正在非常快速的發展，更可預見未來處理問題的範圍可大幅擴大到人類心智所能處的範圍，而後人工智慧即漸漸應用於各行各業上，而人工智慧之演進可分為以下時期：

### 1. 萌發期（1943 年至 1956 年）：

最初的人工智慧研究於 1930 年代末至 1950 年代初，自電腦發明以來，人們便渴望建立能讓電腦擁有類似人類的智慧，此階段的研究重心擺在定理的證明以及通用問題的求解，1950 年圖靈（Turing）發表了一篇論文並預言創造出具有真正智慧機器的可能性，提出如何判定機器具有智慧的試驗「圖靈測試（Turing test）」，並測試機器是否能夠表現出與人類相同或無法區分的智能，圖靈測試（如圖 1-2）至今仍對人工智慧的研究相當重要，而至 1956 年「人工智慧」此名詞才於美國達特茅斯會議中被人工智慧之父 John McCarthy 提出，其討論之問題包含計算機、自然語言處理、神經網路、計算理論以及抽樣化等議題，至今皆成為人工智慧研究發展的重要領域，此會議也成為人工智慧的起源。

### 2. 成長期（1956 年至 1974 年）：

第一波人工智慧自 1956 年被提出後，1957 年由羅森布拉特（Frank Rosenblatt）發明了第一個神經網絡，將人工智慧推向第

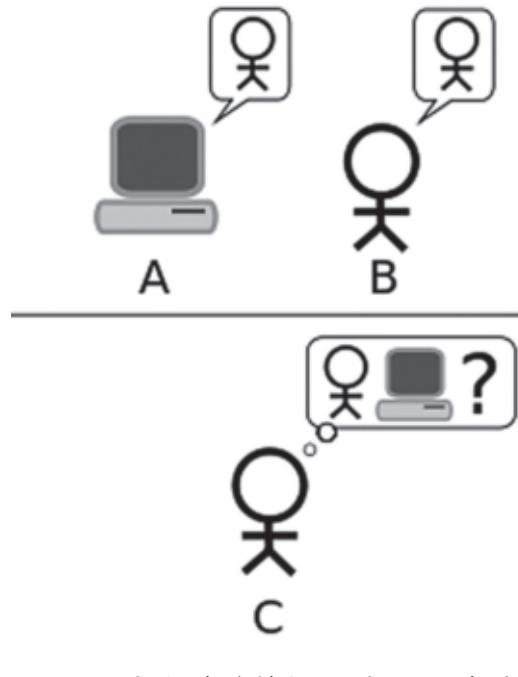


圖 1-2 圖靈測試標準模式：C 使用問題來判斷 A 或 B 是人類還是機械

資料來源：



一次的高峰期，同時，由人工智慧之父 John McCarthy 於 1958 年開發人工智慧語言 LISP 程式語言，機器人學受到重視，此時之研究重點在於使電腦具有理解能力，將人類神經系統與電腦作比較，並研究以機械替代神經系統之可能性。此時期許多知識表示方法問世，例如語意網路（Semantic Network）（是一種有向圖形式；其中，頂點代表的是概念，而邊則表示的是這些概念之間的語意關係，如圖 1-3 所示）及框架理論（Frame Theory）（如圖 1-4），1960 年



## Semantic Networks

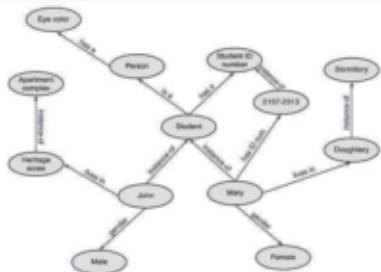


Figure 1-3  
A semantic network

圖 1-3 語意網路（Semantic Network）示意圖

資料來源：



代人工智能的應用皆以「是」與「否」理解，使得應用的面向以代數題以及數學證明為主。由於人工智能的研究方向侷限於邏輯數學領域，加上電腦的運算速度尚未提升、硬體的門檻過低使儲存空間受限制

以及資料量不足夠之困境，因而限制了人工智能的研究與發展，因運算能力沒有突破人工智能，1970 年末知名研究發展計劃如紐厄爾和西蒙「通用問題解決器」以及日本政府領頭的「第五代電腦系統」未得到預期成效時，企業與政府紛紛撤資、研究基金的削減以及計劃的停止，1970 年代進入第一次人工智能發展泡沫。

### 3. 發展期（1974 年至 1993 年）：

1974 年起，因人工智能過了幾年仍沒有重要的大突破，主要受限於計算機內存有限與處理速度慢、應用面較無之前想像中廣闊且無法回答人類不知道的問題，人工智能硬體市場需求的下跌，專業系統的發展下，雖然有許多商業應用的實際案例，但應用範疇亦有限，熱潮也漸漸的衰退，此時亦視為人工智能之第二個低谷。

然而，過去的人工智能判斷問題為真假二元論，1980 年起則添加了量化的概念，描述了問題發生機率，且因電腦技術的發展，使硬體儲存成本下降且運算技術增強，能處理大量數據之分析，1982 年 John Hopfield 發明了霍普蕾爾網路 (HNN)，提出了神經網路將人工智能推向了第二波浪潮，「機器學習」逐漸興起，為人工智能的技術之一，透過演算法使機器能從大量歷史資料中學習規律，並找出資料之特徵與屬性，從而對新的樣本作智慧辨識或對未來進行預測，亦為資料科學

## Relational Frame Theory

A Behavioral Approach to Language

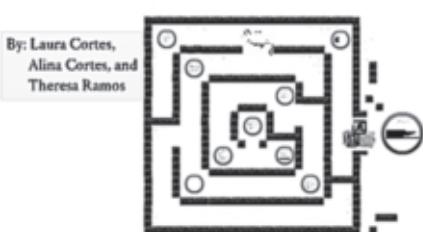


圖 1-4 框架理論（Frame Theory）示意圖

資料來源：



(Data Science) 的熱門技術之一，在 1980 年代至 2006 年間成為研究的主流，其中，機器學習的技術包含支持向量機（Support Vector Machine, SVM）、決策樹（Decision Tree）、羅吉斯迴歸（Logistic Regression）、類神經網路（Artificial Neural Network, ANN）等技術。

在 1986 年 Rumelhart 和 Hinton 等學者提出了反向傳播算法（Back Propagation），解決了神經網路所需要的複雜計算量問題，因而帶動了神經網路的研究熱潮，廣泛的被應用於汽車控制（根據駕駛開車力道與習慣訓練調整機器）、家電控制或是智慧型辨識等領域，然而，研究不久即發現反向傳播的優化（找出誤差的最小值）會發生梯度消失問題（vanishing gradient），使得類神經網路再次凋零，此時的淺層深度學習，如支持向量機（SVM）（如圖 1-5）等技術的興起，能將郵件分為垃圾郵件與非垃圾郵件的功能倍受矚目。

另一方面，1980 年由專家系統之父 Edward Feigenbaum 開發第一套的專業系統（Expert System），用來快速鑑定出化學有機分子，由於人類智慧非常複雜，思考、學習、創造、聯想、類推等能力交錯運用，因此針對特定問題領域全力開發的專家系統，使人工智慧獲得許多令人肯定的結果。這時代的人工智慧透過蘊含統計思維的支持向量機等的機器學習演算法與專

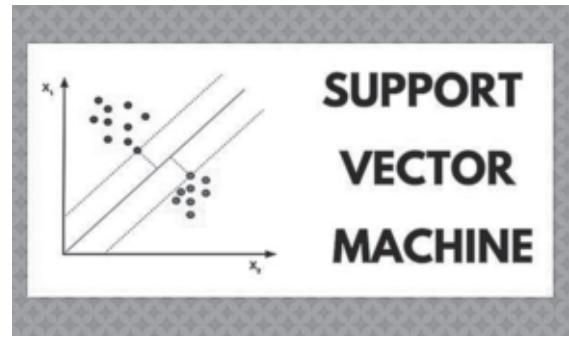


圖 1-5 支持向量機  
(Support Vector Machine, SVM) 示意圖

資料來源：



家系統帶起第二波的熱潮，也把人工智慧漸漸推廣到實用層面。然而，專家系統是由大量的知識庫與推理規則堆疊而成，因此需透過領域專家模擬才能解決教具複雜性的問題，其應用範圍亦仍有侷限性，加上電腦不斷的進步，因此專家系統漸漸走下坡，1987 年起，把人工智慧帶到另一個低點，因人工智慧不能無止盡追逐理想，而必須使其更加實用，否則將成泡沫。

#### 4. 人工智慧發展至今：

從 1993 年起，由於硬體的運算功能大幅提升，機器學習的演算法在語音辨識以及影像辨識的能力也快速發展，其應用領域漸漸的擴大，因此，第三次受到學者與企業的矚目與期望，2006 年 Hinton 提出限制玻爾茲曼機（restricted Boltzmann machine, RBM）模型與深度信年網路（Deep Belief

Network, DBN) 成功訓練多層神經網路，使類神經網絡再度興起，將多層神經網絡 (Deep Neural Network) 稱為深度學習 (Deep Learning)，此為人工智慧發展機會的起源，加上由 Nvidia 成功開發出統一計算架構的整合技術 (CUDA)，透過 CUDA 可撰寫 C 語言使用 GPU 的計算資源加快了深度學習的運算功能。

然而，直至 2012 年才成為人工智慧真正的轉捩點，由 Hinton 的兩名學生透過深度學習技術，利用 GPU 運算，成功讓 GPU 技術被發掘，更使深度學習重回技術熱潮，此時 Google Brain (如圖 1-6) 亦以深度學習從 YouTube 影片中的照片學習辨識貓臉準確度達 75%、Apple 推出了智慧語音助理 Siri (如圖 1-7)，為人工智慧的第三波浪潮。



圖 1-6 Google Brain 運用深度學習技術成功將馬賽克轉換為清晰圖像

資料來源：



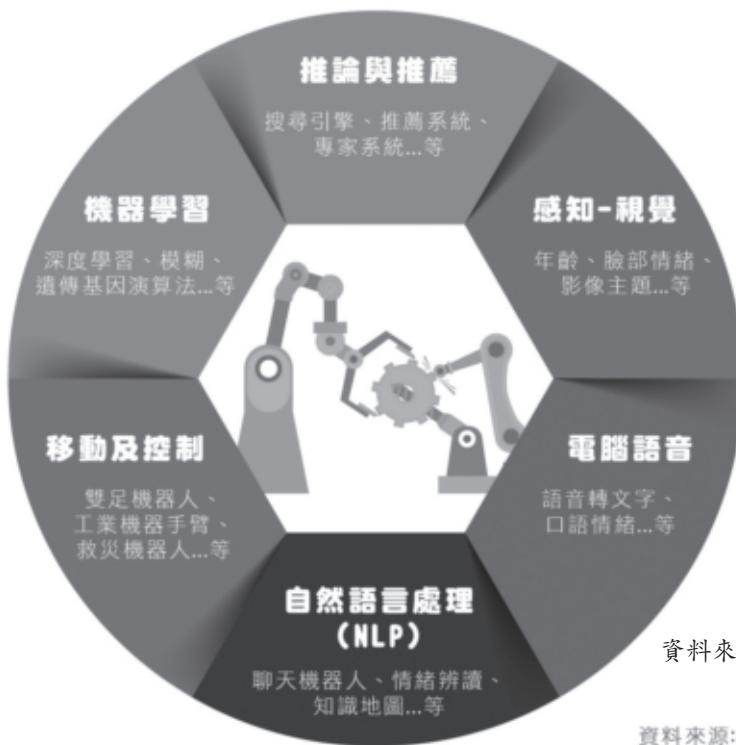
圖 1-7 Apple 公司智慧語音助理 Siri 運用圖

資料來源：  


## (二) 人工智慧 (AI) 的應用

自 2012 年以來，人工智慧開始由研究走入了商用領域，其應用的領域主要分為語音辨識、影像辨識及自然語言處理等三部分，其中語音辨識至今已有等同於人類的辨識度，其主要的科技巨頭 Apple、Google 及 Amazon 也陸續提出應用於日常生活的服務；第二領域則為影像辨識，從靜態圖片辨識到動態影像的辨識，雖然動態影像辨識準確度仍跟不上人類，但最受注目的應用即為自動駕駛的技術，Google

## 人工智慧主要技術分類



資料來源：資策會



資料來源：資策會MIC，2017年6月

圖 1-8 人工智慧主要技術分類與相關應用

持續在進行自動駕駛的研究，且 TOYOTA 也在美國設立研究所，目前影像辨識仍在研究至實用的過程；最後為自然語言處理（Natural Language Processing, NLP），目的為讓人工智慧理解人類所表達的語意，從分解詞性的「語素分析」（Morphemic analysis），分解出最小的字義單位後，再進行「語法分析」（syntactic analysis），最後再透過「語意分析」（Semantic analysis）來了解含意。

由於網路的快速普及，高性能電腦、大數據探勘（Data Mining）及機器學習（Machine Learning）的持續發展與進步，使人工智慧在近幾年有更大的技術突破，更興起對機器學習技術的應用，其中，2010 年人工智慧浪潮中興起的機器學習技術為類神經網路（Artificial Neural Network, ANN），屬於深度學習（Deep Learning）之技術之一，透過模仿生物神經網路的結構與功能所產生的數學模型，2014 年

Facebook 推出了臉部辨識系統 DeepFace 準確率高達 97%、2015 年 Google 發表開源深度學習系統 TensorFlow 且微軟亦推出了個人數位語音助理 Cortana，2016 年大家耳熟能詳的事蹟即為 Google DeepMind 的 AlphaGo 打敗了韓國圍棋棋王、Facebook 推出了「Facebook Messenger Platform」等技術，使 NLP 技術的聊天機器人更受矚目，其應用領域相當廣闊，包含金融科技、智慧製造、醫療健康、交通方面等，而自動駕駛與智慧機器人將會是第一波實現人工智慧社會之重點。

人工智慧發展至今以相當快的速度在成長，從 2000 年起迄今，人工智慧領域的投資規模相當龐大，美國人工智慧行業的創業公司數量大幅增加約 14 倍，且風險投資對此領域投資金額更成長至 6 倍（如圖 1-9 與圖 1-10 所示），以台灣為例，除了簡化醫療保健行業的工作流程、減少了教育部門行政任務的工時外，同時也降低了製造業的管理成本。Adobe（2018）於 2018 年數位趨勢報告中，全球最佳企業利用人工智慧做市場行銷的比例，高於其他企業兩倍以上（分別為 28% 及 12%），相較北美及歐洲，亞太區受訪者更傾向使用整合、以雲端為基礎的科技，此顯示了亞太區市場行銷人員更擅長運用新科技以傳遞個人化內容，使工作更加快速且智慧化，未來提升客戶體驗正是市場行銷人員 2018

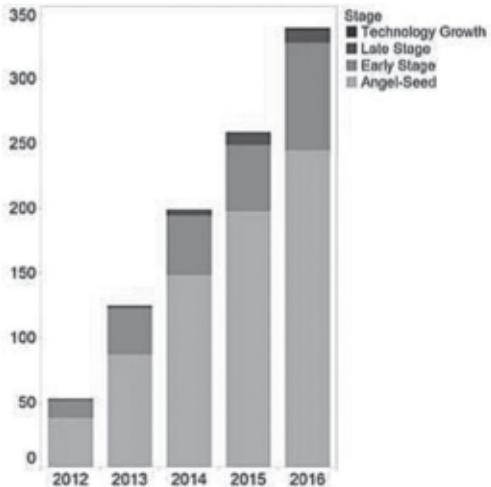
年後的首要任務為提升顧客體驗，將購物體驗加入個人化以及顧客之關聯性，提供客製化之商品與服務，市場研究中心 Gartner 也在 2018 年的研究報告中提出人工智慧行業的總價值至 2018 年高達 1.2 萬億美元之高，相較於 2017 年成長大約 70%，更預測未來人工智慧的商業價值將高達 3.9 萬億美元，在此計算能力、容量、速度和各種數據的進步的世代中，加上深度學習（DNN）的快速發展，人工智慧為未來的 10 年最具有顛覆性的技術，其中，創造商業價值最大的領域極為顧客體驗之解決方案。

自 2017 年 AlphaGo 開啓了人工智慧應用的發展歷程，而後 2018 年 google 與 Microsoft 都相繼把內部的機器學習資源解密，深度學習的門檻會大幅降低，因此，企業下一步的重點應在於資料洪流中取得正確的資料以及大幅的應用，深度學習演算法發展成功是個起點，要能落實到各行各業，使得產業人工智慧化，同時亦需要從各產業內部先盤點需求與痛點，透過演算法把所有的產品融合最終為一個解決方案，才能達到產業升級的終極目標。

## 貳、大數據分析應用於零售商

在我們生活中，零售店面無所不在，不論是便利商店、百貨公司、商品專賣

## **artificial intelligence funding rounds**



## **artificial intelligence invested amounts**

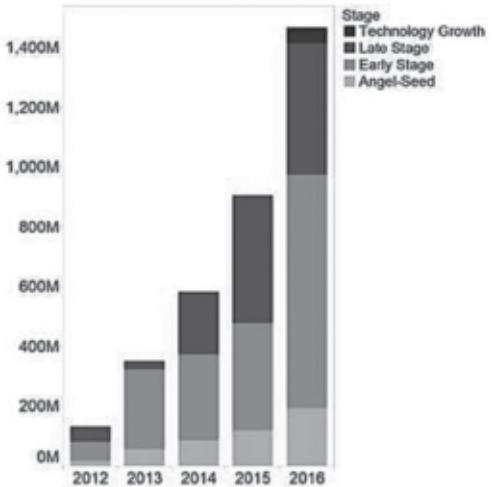


圖 1-9 人工智慧的投資  
資料來源：crunchbase1 annual 2017



### **artificial intelligence invested amounts heatmap**



圖 1-10 全球人工智慧投資熱度地圖

資料來源：crunchbase1\_annual\_2017



店、購物中心等都屬於零售業的一部分，也都擁有各自的店面、門市。由於現今網際網路的快速發展、互聯網的興盛、智慧



資料來源：

行動裝置的普及，加上大數據時代的來臨，技術不斷的推陳出新，不僅改變了實體店面在電子商務環境中所扮演的角色和功能，同時也改變一般人們的購物模式。依據 Laudon 對電子商務的描述，互聯網商業又被稱為電子商務，也就是任何形式的商業交易或使用 ICT (Information Communication Technology) 進行訊息交換的模式，都稱之為電子商務。電子商務將商品從原本的實體店面，改為在網路平台販售，讓商品與網路結合。另外伴隨著儲存技術的發展，電子商務除了能降低門市所消耗的營運成本、流動成本外，還能容納大量的商品和儲存更多的商品相關資訊。但網路購物雖然方便，卻無法像實體店面一樣能真實的觸摸和觀看實體物品。因此在電子商務成熟的年代，零售商的實體店面如何與之競爭，轉而更需要注重的是顧客體驗與各品牌互動的滿意程度以及如何創新、整合和自動化，這將影響未來零售業在實體店面的營運模式。

「Big Data」最早出現在 2010 年，由 IBM 電腦公司所提出，翻譯成中文則有「海量資料」、「巨量資料」等多種說法，目前相關文獻則是以「大數據」為統稱。在 2001 年，Gartner 公司的分析家 Doug Laney 首先發表了一篇研究報告指出資料處理可分為三個所需面對挑戰和機會的關鍵：資料量 (Volume)、速度 (Velocity) 和多

樣性 (Variety)。在 2012 年，Doug Laney 更進一步的定義了大數據：「大數據是大量、高速以及類型多變的資訊資產，它需要全新的處理方式，去促成更強的決策能力、洞察力與最佳化處理。」除了以上所敘述的三個關鍵外，Inderpal Bhandar 認為應該加入真實性 (Veracity)，分析過濾資料的不完整、不一致、假資料等情況發生，避免影響損害資料系統的完整性和正確性，另外合法性 (Validity)、可視性 (Visualization) 和價值 (Value) 也逐漸被討論。現今在商業的競爭中，數據分析已成為輔助營運的一種有效手段，因此利用 Hadoop 建構電子商務分析平台分析和處理電子商務數據，以便讓企業更清楚了解客戶的偏好，從而為消費者提供更有針對性的產品或服務，以滿足客戶的實際需求。不尋常的數據實例被稱為不平衡數據，因此提出一種處理不平衡數據的方式，希望能提高數據的可用性，這些數據集用於個性化電子商務的客戶體驗，在預測準確性、精確度和特異性之間達到最佳平衡。淘寶網為亞洲最大的零售購物網站，因此針對阿里巴巴的淘寶網電子商務做了一系列的研究，發現在 2008 年阿里巴巴集團就意識到了大數據的重要性，將之置於戰略決策之一，從信息技術 (IT) 到數據技術 (DT) 都有著明確的未來發展戰略。研究將大數據應用於跨境電子商務的準確營

銷，可以得到的結論為大數據不僅能為消費者提供便利外，還可以為企業帶來多方面的好處。

### 一、大數據分析應用於零售商推薦系統

隨著數位傳播技術的快速變遷，現今的資訊量更是難以數計，在資訊傳遞出推薦系統的應用。推薦系統是指能夠幫助消費者即時找到與其喜好相關的商品資訊，而如今推薦系統已經接續成為各行各業所應用的系統，其應用包含：音樂、旅遊、書籍、電影、衣服…等等，著名的推薦系統像是 Amazon (O'brien & Marakas, 1994)，這家公司能根據讀者的習性及偏好等推薦不同的產品資訊給消費者，如此一來，不但能節省顧客搜尋產品的時間，與沒有導入推薦系統的公司相比，能吸引更多消費者前往購買產品。

#### (一) 推薦系統 (Recommendation Systems)

推薦系統是一個能夠推薦並提供消費者購買商品的參考依據。這些建議則是根據許多的決策而來，例如消費者買了什麼產品？消費者看了哪部電影？或是消費者在線上閱讀了什麼文章？(Ricci, Rokach, & Shapira, 2011)

由於數位資訊的爆炸性成長以及大量使用網路的訪客增加，使得資訊超載成為了現今潛在的挑戰，人們想要即時在網路上取得感興趣的資訊，這也是推薦系統的

需求增加的主要原因 (Isinkaye, Folajimi, & Ojokoh, 2015)。而推薦系統能依據使用者的喜好與興趣，篩選出重要且對使用者有用的資訊，因此，推薦系統可以解決現今資訊超載的問題。另外，推薦系統也能夠針對特定的使用者，根據與該使用者有類似偏好的其他使用者，預測其可能會感興趣的產品 (Isinkaye et al., 2015)。也就是推薦系統常見的內容過濾 (Content-based filtering) 及 協同過濾 (collaborative filtering) 方法。推薦系統對於使用者，可以大幅縮短他們瀏覽大量資訊的時間並能快速地挑選適合自己的產品；對於服務提供者，導入推薦系統能夠幫助他們的顧客即時找到感興趣的產品，如此一來就會有更多消費者願意在該服務平台上購買產品，而成為了忠誠的顧客。

推薦系統的運作週期如圖 2-1 所示，首



圖 2-1 推薦系統運作週期模式  
資料來源：Isinkaye et al. (2015) 及本文整理

先收集使用者的資訊，包含喜好及購買過的產品等，接著系統會自主學習並建立模型，最後預測使用者可能會感興趣的產品並推薦之，而系統則會收集使用者選取的資料並回到第一階段反覆執行（Isinkaye et al., 2015）。

目前常見的推薦系統方法有以下幾種：

#### 1. 非個人化推薦（Non-Personalized Recommendations）：

非個人化推薦是將其他所有使用者對於產品的評價取平均值，而後將這些平均值排序後依依推薦給使用者，因此這類的

推薦方法會給每位使用者有相同的推薦，類似「暢銷產品」的概念，將較受喜愛的前幾項產品推薦給消費者（Schafer et al., 1999）。

#### 2. 內容過濾（Content-based filtering）：

內容過濾的推薦系統是根據使用者歷史消費紀錄或過去喜愛的產品來學習，系統會推薦與歷史資料相似的產品給使用者。例如，使用者過去看過屬於喜劇的電影，那麼推薦系統就會推薦其他同樣屬於喜劇的電影給使用者（Ricci et al., 2011）。圖 2-2 為內容過濾運作的流程圖，本文將其

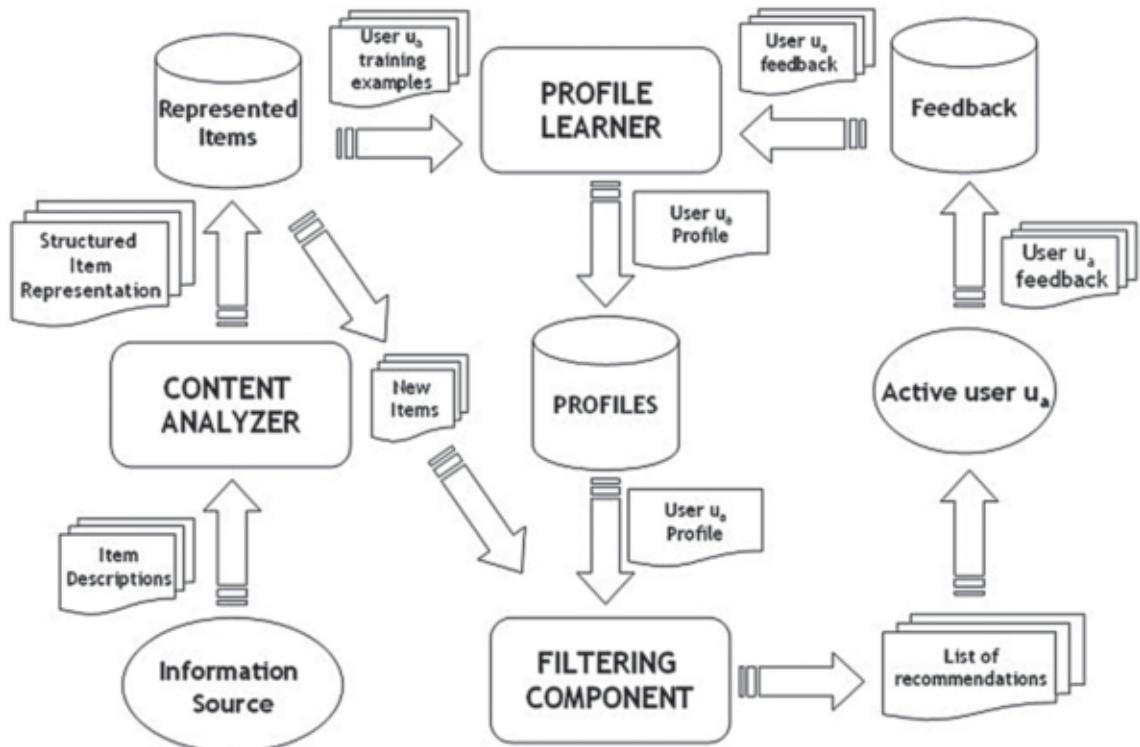


圖 2-2 內容過濾推薦法的高階架構圖

資料來源：Ricci et al. (2011)

流程分為三個部分，分述如下：

(1) 從新項目取得資訊：當有新的產品項目時，擷取項目的特徵描述，進行內容分析，若有非結構化的資料（例如：文字檔），轉成結構化項目（例如：產品描述向量）呈現，儲存至項目的資料庫中，當作使用者檔案和內容過濾法的輸入。

(2) 從使用者檔案取得資訊：根據使用者的歷史消費記錄與喜好項目記錄，存入檔案的資料庫中，當作內容過濾法的輸入。

(3) 從推薦項目取得資訊：使用者點選推薦系統所推薦的項目，也會回饋到使用者歷史消費記錄或喜好項目中，如此循環內容過濾的推薦。

### 3. 協同過濾 (Collaborative filtering) :

協同過濾是根據與使用者有相似興趣的其他使用者，將其他使用者過去的喜好推薦給該使用者。兩者的相似性則是由每位使用者過去對項目的評分，來計算使用者間的相似度 (Ricci et al., 2011)。而協同過濾又分為以「使用者」為基礎的協同過濾與以「項目」為基礎的協同過濾，這部份會在下一節進一步探討。

### 4. 混 合 式 推 薦 系 統 (Hybrid recommender systems) :

所謂混合式推薦法是指將上述兩種方法（內容過濾和協同過濾）結合的技術，透過內容過濾的優點補齊協同過濾的缺點；

反之，透過協同過濾的優點補齊內容過濾的缺點，例如，協同過濾法會受新項目的限制，在所有使用者皆未對新項目評分的情況下，協同過濾法就無法對使用者進行推薦；但由於內容過濾是針對項目的描述（特徵）進行推薦，而項目的描述則是容易取得的，因此內容過濾並不會受新項目的限制 (Ricci et al., 2011)。而推薦系統除了推薦的方法之外，使用者對項目的評分資料也分為隱式資料 (Implicit Data) 與顯式資料 (Explicit Data) (Breese, Heckerman, & Kadie, 1998)，顯式資料指的是使用者能明確對項目的喜好評分，例如 1 到 5 分的衡量尺度；隱式資料則是指使用者的瀏覽紀錄、消費歷史紀錄等無法明確表明使用者的喜好程度之資料。而本文無法明確得知消費者對每個購買產品的評分，也就是消費者對購買產品符合其喜好的程度如何，因此，採取蒐集隱式資料的方式，將消費者購買產品的次數當作使用者對該項目的評分。

### (二) 協同過濾 (Collaborative filtering)

協同過濾旨在找出與目標使用者有相似喜好的其他使用者，而 Schafer et al. (1999) 認為「使用者間相關性 (people-to-people correlation)」的推薦方法指的是使用者在電子商務網站上購買商品的相關性。圖 2-3 為協同過濾運作的流程圖，其流程步驟如下：

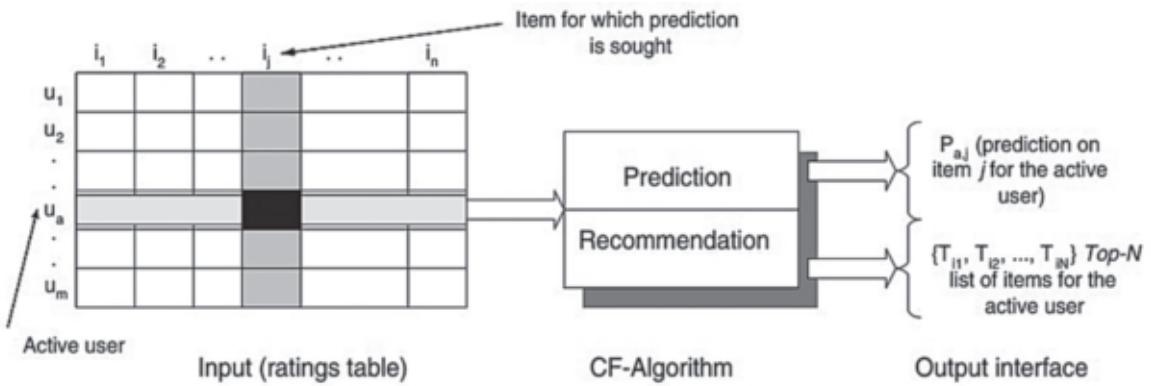


圖 2-3 協同過濾流程圖

資料來源：Sarwar, Karypis, Konstan, and Riedl (2001)

1. User-item matrix：將使用者跟產品項目建立成一個矩陣（user-item matrix），記錄使用者對項目的評分。

2. CF Model：利用相似度演算法預測使用者對項目的評分，例如：最近鄰居法（K-Nearest Neighbor, K-NN）（Punch III et al., 1993）。

3. Prediction：將使用者尚未評分的項目，根據與使用者有相似興趣的其他使用者對該項目的評分，當作使用者對該項目的評分。

4. Recommendation：對所有使用者感興趣項目進行預測，最後，利用 Top N 推薦產生推薦結果，找出與使用者近似的群體評分較高且使用者尚未消費或評分的項目，推薦給使用者。

協同過濾的技術最早是應用於電子郵件的過濾上，帕羅奧多研究中心（Palo Alto Research Center）開發了一套 Tapestry 系

統，而 Goldberg, Nichols, Oki, and Terry (1992) 提出協同過濾的方法，解釋此系統可以儲存過去所傳送過的電子郵件，透過使用者感興趣或不感興趣的郵件議題，進而過濾所有電子郵件給使用者，讓使用者不需要花費大量的時間來搜尋他想看的電子郵件，但這套系統需要透過使用者輸入欲搜尋的關鍵字來過濾電子郵件，而 Resnick, Iacovou, Suchak, Bergstrom, and Riedl (1994) 開發出一套 GroupLens 系統，此系統為採取匿名或非匿名的方式，蒐集使用者對網路新聞的評分（最高 5 分，最低 1 分），根據這些評分並利用協同過濾的方法，預測其他使用者對該篇網路新聞的喜爱程度，進而推薦使用者其可能喜歡的網路新聞，現在有許多電子商務網站都利用自動過濾推薦的方式，推薦消費者可能會感興趣的商品，例如：Amazon、Yahoo 超級商城等等（如圖 2-4 與圖 2-5 所示）。

## Neighborhood based collaborative filtering

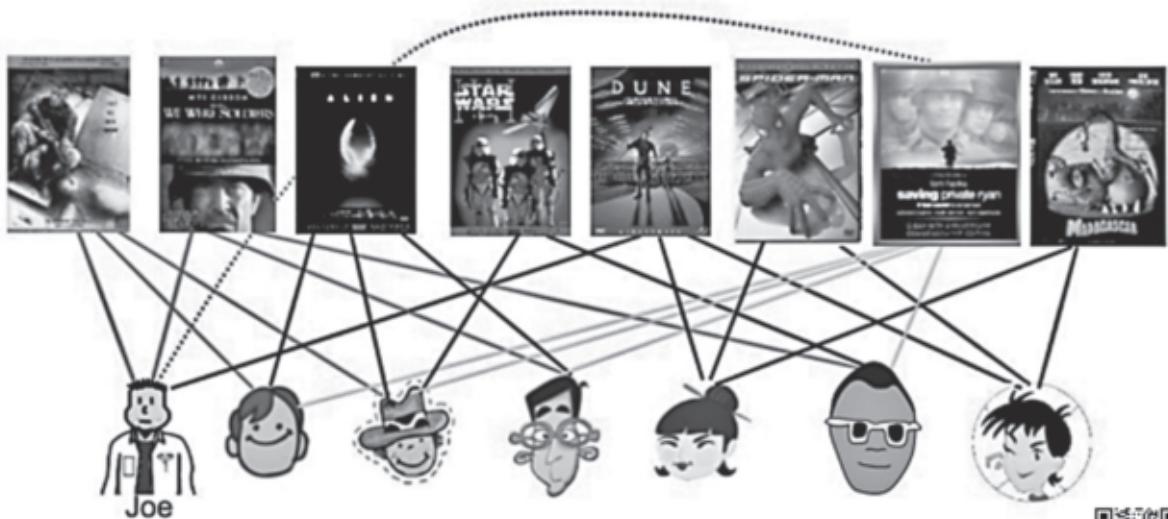


圖 2-4 協同過濾運用於影片選擇分析圖示

資料來源：



The screenshot shows the Yahoo Super Mall homepage. The search bar at the top contains the text '請輸入商品或商店關鍵字' (Please enter product or store keyword). Below the search bar are navigation links for '商品分類' (Product Category), '逛商店' (Browse Stores), '搜尋商店' (Search Store), and various promotional banners for '漢神百貨' (Hanshin Department Store), '流行館' (Pop Culture), '型男館' (Men's Style), '購划算' (Buy Smart), '聯名卡' (Co-branded Card), and '超贈點' (Super Points). A banner for '德瑞健康家' (Drei Health Home) is also visible.

The main content area features a large advertisement for 'PHILIPS FHD 液晶顯示器' (PHILIPS FHD LCD Monitor). It highlights a promotion where a 43-inch monitor is offered for \$888. The ad includes a woman smiling and various model options. Below this, there are sections for '今日爆殺，一日限定' (Today's Kill, One Day Limited) and '團購風暴，低價來襲' (Group Purchase Storm, Low Price襲來). A banner at the bottom left reads '今日爆殺限定跨店滿減好好買' (Today's Kill Limited across stores, full reduction, buy well).

圖 2-5 Yahoo 超級商城透過協同過濾法展示近期查詢商品

資料來源：



協同過濾主要的技術分為兩種（如圖 2-8 所示），「以使用者為基礎的協同過濾」與「以項目為基礎的協同過濾」，以下做進一步的探討。

### 1. 以使用者為基礎的協同過濾

以使用者為基礎的協同過濾（User-based Collaborative Filtering），最早由 Schafer et al. (1999) 提出，指的是依據使用者之間的喜好相似程度進行推薦，例如以某消費者與其他消費者在電子商務網站上購買商品的關聯性來推薦其可能會喜歡的商品。其演算法是利用所有 User 和 Item 的資料庫去預測 User 對 Item 的評分，最常用的技術是最近鄰居法，找出與使用者對相似的項目進行評分且皆為類似評分者，即為該使用者的鄰居，接著透過鄰居所評分的其他項目對該使用者對這些項目進行預測，且利用 Top-N 的推薦法推薦較感興趣的前 N 名項目（Sarwar et al., 2001）。

### 2. 以項目為基礎的協同過濾

以項目為基礎的協同過濾（Item-based Collaborative Filtering），最早由 Schafer et al. (1999) 提出，指的是依據使用者感興趣的項目進行推薦，例如某消費者將一些商品放置購物車中，推薦系統將會推薦與這些商品相似的其他商品給消費者，進而擴大消費者的購買商品數量。其演算法是計算兩個項目間的相似性，找到與目標項目相似的其餘項目，首先，必須從不同使

用者對項目的評分矩陣 ( $m \times n$ ) 中找出同時有被使用者評分過的項目 i 與項目 j (如圖 2-6)，進而去計算兩者間的相似性，例如 Cosine similarity，接著透過加權的方式，對與目標項目相似的其他項目進行排序，進而預測使用者對這些項目的評分（如圖 2-7），且利用 Top-N 的推薦法推薦較感興趣的前 N 名項目。

## （三）情境感知推薦（Context-Aware Recommendation）

傳統推薦系統的資料維度為 2，也就是使用者與項目的評分矩陣，而情境感知推薦（Context-Aware Recommendation）是將評分矩陣變成多維度（Multi-dimensional）的概念，也就是使用者、項目與情境（Context）(Adomavicius & Tuzhilin, 2011)，指的在推薦時，加入使用者給予評分時的情境或是消費者所設定之推薦情境進行推薦，包含人物、時間、地點、事件等，例如，在推薦電影的系統中，能夠根據使用者看電影的時間進行推薦，若某使用者喜歡在周末看電影，那麼，系統能篩選出所有在周末看電影之使用者的評分中，平均評分最高的電影給這位使用者參考；另外，使用者在進行推薦系統的研究時，也能夠自行設定推薦的情境，從不同的情境研究推薦的效能。在情境感知推薦系統中，可想而知，情境資料的蒐集來源為我們需要注重的環節，如同第一節所

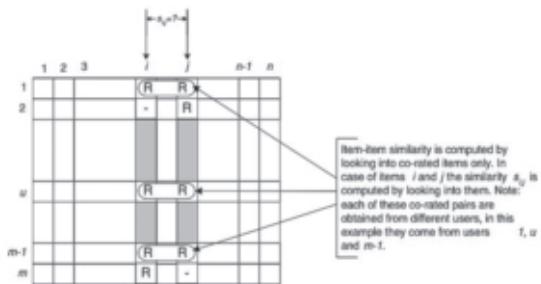


圖 2-6 共同被評分的項目並計算相似性之示意圖  
資料來源：Sarwar et al. (2001)

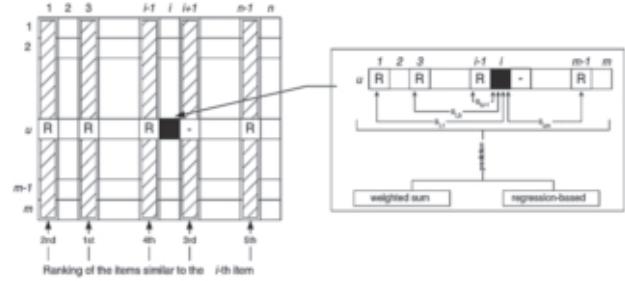


圖 2-7 以項目為基礎的協同過濾演算法之預測過程  
資料來源：Sarwar et al. (2001)

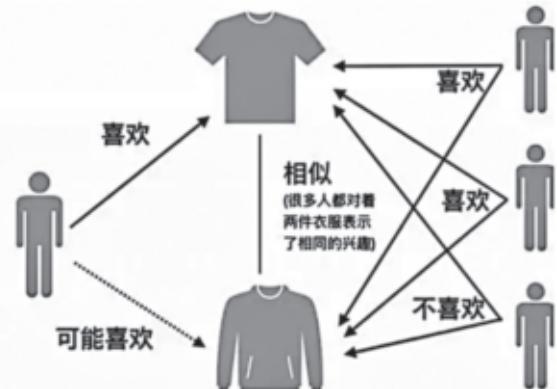
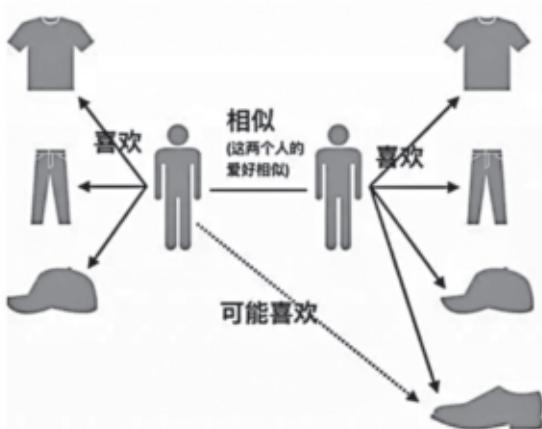


圖 2-8 協同過濾主要技術圖示

資料來源：



述，推薦系統之資料可分為隱式資料 (Implicit Data) 與顯式資料 (Explicit Data)。此節的隱式資料為使用者的瀏覽商品或看電影的時間等，使用者不必明確指出看電影的時間，系統便能透過網頁瀏覽的記錄，將時間蒐集至情境感知推薦系統的研究中，而這部分的資料大多經由提供線上觀看電影的網站取得；顯式資料則是使用者直接標明看電影的時間、與誰一同

觀看電影或看電影的地點，但對於顯式資料的取得，可能會有隱私的問題產生 (Liu, Zhang, Hui, & He, 2015)，使用者對於觀看時間、與誰觀看以及觀看地點可能會有隱瞞之行為，對於不想透露太多個人隱私的使用者，研究人員在蒐集情境資料上便會有相當的難度，因此，情境感知推薦系統的資料選取也是一門重要的課題。

此外，Adomavicius and Tuzhilin (2011)

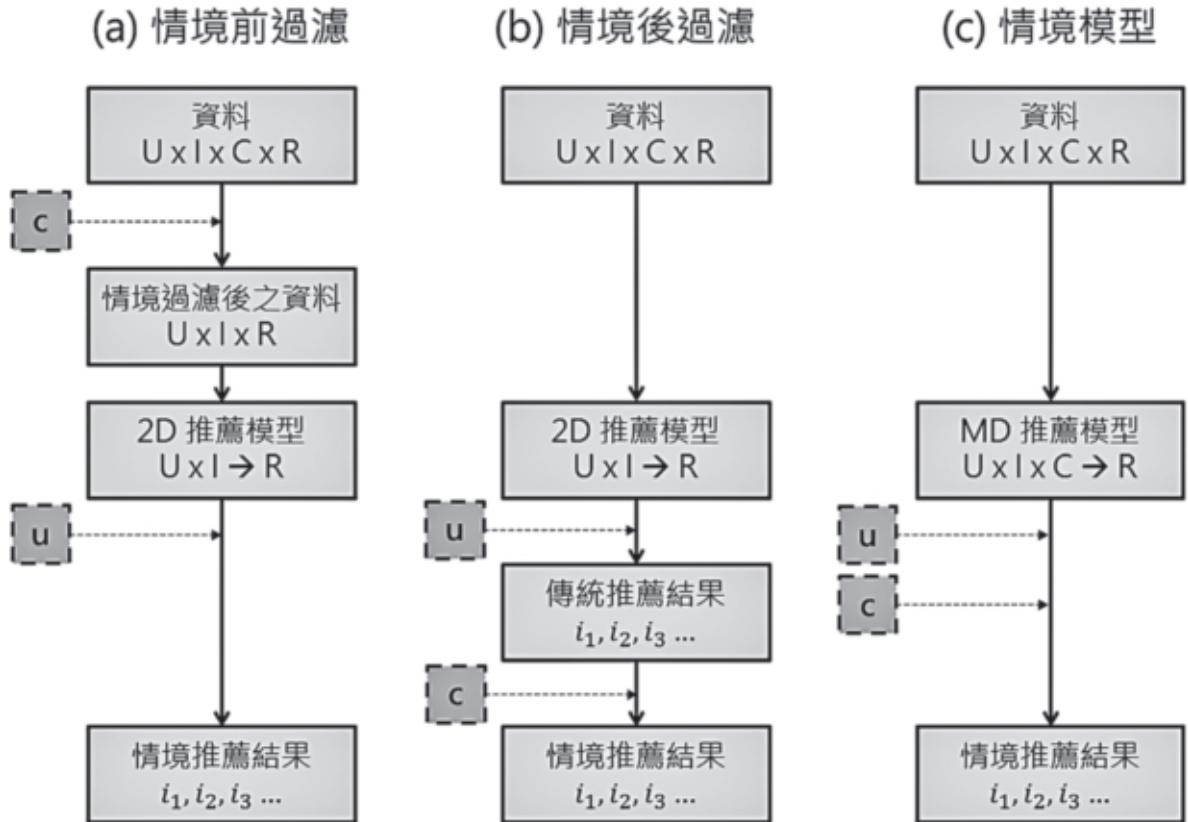


圖 2-9 情境感知推薦之示意圖

資料來源：Adomavicius and Tuzhilin (2011) 與本文整理

提出情境感知的推薦分為三種：(1) 情境前過濾（Contextual Prefiltering）：在進行推薦前，先根據所設定之情境過濾原始資料（如圖 2-9 (a)）；(2) 情境後過濾（Contextual Postfiltering）：在進行推薦後，根據所設定之情境過濾推薦資料（如圖 2-9 (b)）；(3) 情境模型（Contextual Modeling）：在推薦模型中加入情境資料進行推薦（如圖 2-9 (c)）。

#### (四) Spark

Apache Spark 是一個叢集式運算框架，由加州大學柏克萊分校的 AMPLab 所開發，而這個模型框架是由 MapReduce (Dean & Ghemawat, 2008) 演變而來，提供資料串流的處理，支援與 MapReduce 相似的特性，包含可擴展性及容錯性，除此之外，Spark 在執行速度上比 Hadoop 快 10 倍，而 Spark 的核心是彈性分散式資料集 (Resilient Distributed Dataset, RDD)，由

Zaharia, Chowdhury, Franklin, Shenker, and Stoica (2010) 所提出的抽象概念，為一種分散式的記憶體，使用者可以從記憶體中存取檔案的地方（例如資料庫）產生 RDD，進行類似 MapReduce 的平行運算，例如資料的轉換與處理，而 RDD 有相當高的容錯性，若其中一部份的 RDD 遺失了，也能有足夠的資訊，再從其他 RDD 中重新轉換產生一個 RDD，而這些 RDD 只能讀取不能修改。

RDD 主要的平行運算分為兩種，轉換 (Transformation) 與行動 (Action) 的操作，轉換的操作是對已存在的 RDD 進行轉換，例如 map、filter、join 等，這些轉換會產生新的 RDD，而轉換的運作為 lazy load 機制，在轉換的動作不會立刻執行，直到

行動操作時才會真正被執行，例如 reduce、collect、count 等。Lazy load 機制為 Spark 在處理資料上的核心，該機制考慮到記憶體的存取空間，倘若每次轉換後所得到的 RDD 都馬上執行並儲存於記憶體，等到行動操作時才來使用，會占用許多的儲存空間，因此，此機制能解決大量的數據在資料處理及轉換後，需要龐大的儲存空間的問題，如此，不必每次轉換都占用一些的儲存空間，可以等到進行 Action (行動) 操作時，再選擇需要被執行的轉換動作。本文將轉換與行動常用的指令整理如表 2-1 及表 2-2 所示：

## 二、小結

歸納整理上述內容，本文主要探討利

表 2-1 spark 轉換 (Transformation) 指令

轉換	指令說明
map(func)	將 RDD 中的元素透過 function 運算產生新的 RDD，例如：將所有元素加 1
filter(func)	將 RDD 中的元素透過 function 進行篩選產生新的 RDD
flatMap(func)	類似 map()，但每個元素的輸出可能是零或多個項目，例如：(Hi, spark) → H, i, s, p, a, r, k
intersection(otherDataset)	對 RDD 進行交集運算
distinct([numTasks])	去掉 RDD 中重複的資料
groupByKey([numTasks])	合併 RDD 中的元素，例如：針對某一欄位的資料進行數值的加總或取平均值
join(otherDataset, [numTasks])	對 RDD 進行關聯合併

資料來源：“Spark Programming Guide”



表 2-2 spark 行動（Action）指令

行動	指令說明
<code>reduce(func)</code>	將存在的 RDD 中所有元素透過 function 進行相對應的運算，例如： $(x, y) \Rightarrow x + y$
<code>collect()</code>	回傳所有進行運算或篩選後的 RDD
<code>count()</code>	回傳 RDD 的元素數
<code>first()</code>	回傳 RDD 中第一個元素
<code>take(n)</code>	回傳 RDD 中第 n 個元素
<code>foreach(func)</code>	將 function 套用到 RDD 的每個元素中
<code>save(path : String)</code>	將 RDD 輸出到檔案分享的系統，例如：HDFS(Hadoop Distributed File System)

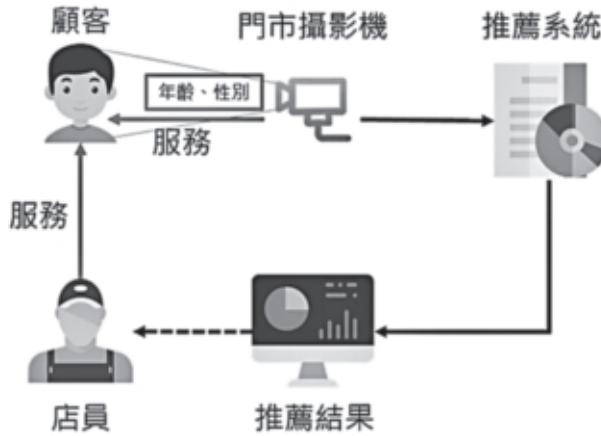
資料來源：“Spark Programming Guide”

用用戶資料中紀錄的屬性，做為用戶彼此間相似度的依據，經過相似度的篩選後，利用權重的方式生成用戶對物品的評分，產生推薦的結果，藉此在傳統的協同過濾方法中加入沒考量到的因素以及增加推薦的精確度。門市業目前最大的問題在於販售商品的方式多為被動的行銷，都是由客戶到店面詢問或者經由行銷廣告，才會參考口頭或紙本所推薦的東西做購買。但現今社會競爭越來越激烈，這樣的銷售手法無法有效吸引客戶進行購買。因此我們利用消費者的會員資訊，每當消費者到門市消費後，系統便會記錄該客戶所購買商品行為等相關資訊，並將門市過往的交易資料套用到推薦系統中，接著我們透過交易紀錄分析客戶的歷史消費行為，推薦其他不同且客戶可能感興趣的商品（如圖 2-10

所示），以增加更多商品的有效銷售。另外我們還能透過歷史資料，分析每個客戶消費的行為，並依據購物籃分析和關聯規則適時推薦其他商品給其他的客戶購買。這些商品很可能是門市員工對顧客介紹時，無法立即想到或忽略的商品。因此我們透過分析資料庫中消費者購買商品的紀錄，找出商品與商品間的關聯性，同時根據用戶過往的交易紀錄結合其他用戶的相似決策建立模型。由此我們可以知道該消費者除了來店裡購買的商品之外還有可能對其他什麼東西感興趣，進而推薦消費者購買。

## 參、深度學習（Deep Learning）

深度學習是機器學習的一個分支，是一種基於對資料進行表徵學習的演算法，



資料來源：



圖 2-10 零售門市推薦系統圖示

深度學習目前已有數種框架，如：卷積神經網絡與遞迴神經網絡，深度學習的應用領域相當廣泛，如：圖像識別、語音識別以及自然語言處理。

### 一、深度學習的定義

深度學習是機器學習的一種方法，機器學習技術就是讓機器可以自我學習的技術，一般機器學習方法要經過三個步驟：首先，人類提供機器一個由函數構成的集合，接著，人類根據訓練的資料定義函數的優劣，最後，機器自動從函數集內找出最佳的函數。

深度學習亦是如此，深度學習是讓機器模擬人腦的運作方式，進而和人類一樣具備學習的能力，它讓電腦可以自行分析資料找出「特徵值」，而不是由人類來決定特徵值，就好像電腦可以有「深度」的

「學習」一樣。深度學習的三個步驟和機器學習一樣，包含：選擇函數集、定義優劣以及挑選函數，其中，人類提供的函數集是由神經網絡的結構所定義，神經網絡和人腦有幾分相似之處，人腦是由神經元所構成，神經網絡也是由神經元連接而成。

深度學習使用多層神經網絡，理論上隱藏層越多自由度與精確度就越高，但是結果剛好相反，因為誤差反向傳播很難傳遞回更上一層的神經元，當階層太多時效果不佳，因此深度學習不但使用多層神經網絡，同時還使用自動編碼器來進行非監督式學習（李宏毅，2016；曲建仲，2017）。

### 二、深度學習的發展

人工智慧、機器學習以及深度學習是密切相關的，深度學習是機器學習的分

支，機器學習又為人工智慧的分支。人工智慧在 1950 年代的發展是以數理邏輯為基礎。機器學習是人工智慧領域的一個重要學科，1980 年代多層神經網絡失敗，淺層機器學習方法興起，此時，機器學習在演算法、理論以及應用等方面都獲得巨大的成功。深度學習是機器學習的第二次浪潮，2006 年 Hinton 成功訓練多層神經網絡，Hinton 和他的學生在 Science 期刊上發表了一篇文章，讓深度學習在學術界和工業界獲得了廣大的關注，但深度學習真正的轉折點是在 2012 年的 ImageNet 比賽，這次的比賽讓深度學習與圖形處理器（Graphics Processing Unit，簡稱：GPU，如圖 3-1）一戰成名，同時，NVIDIA GPU 也成為運算中不可或缺的硬體，直到現在，深度學習已經成為大數據和人工智慧的一個熱潮（余凱、賈磊、陳雨強、徐伟，2013；Lynn, 2017）。

深度學習的概念源自於神經網絡的研究，神經網絡技術在 1980 年代就已經發展成熟，只是當時並不常用「深度學習」這個詞彙（李宏毅，2016）。直至今日，在深度學習領域中，神經網絡仍是相當重要的核心價值，神經網絡的發展深深影響著深度學習，Schmidhuber（2015）針對神經網絡的發展進行了詳細的回顧，本文將其整理如表 3-1 所示。

隨著深度學習技術之發展，目前已經



圖 3-1 NVIDIA 圖形處理器

資料來源：



有許多針對不同資料特性的類神經網路架構，有些適用於圖像的辨識、分類；有些則適合針對有前後關係的時間序列資料進行處理，例如文章、語音的辨識，以下將針對神經網路架構做簡單介紹。

#### （一）多層感知器運算模式

多層感知器（Multilayer Perceptron, MLP）是最常見的類神經網路架構之一，在其中具有隱藏層，使其具備解決非線性分類問題的能力，其基本架構如圖 3-2 所示，分為輸入層（Input Layer）、輸出層（Output Layer）以及隱藏層（Hidden Layer）。前一層神經網路的輸出訊息經由輸入層接收後，傳遞至中間的隱藏層。隱

表 3-1 神經網絡的發展

年代	發展
1940 年代	早期的神經網絡，還沒有學習。
1960 年左右	視覺皮層為深度學習提供了靈感。
1960~1981 年	為神經網絡開發反向傳播（Backpropagation，簡稱：BP）。
1965 年	基於數據處理分組方法（Group Method of Data Handling，簡稱：GMDH）的深度網絡。
1979 年	引入卷積神經網絡（Convolutional Neural Network，簡稱：CNN）、權重重置以及色度抽樣。
1980~2000 年	神經網絡的大量改進。
1987 年	通過自動編碼器（Autoencoder，簡稱：AE）層級的非監督式學習（Unsupervised Learning，簡稱：UL）。
1989 年	卷積神經網絡的反向傳播。
1991 年	基礎深度學習梯度下降的問題與時間遞迴神經網絡（Recurrent Neural Network，簡稱：RNN）。
1992 年	卷積神經網絡的最大池化（Max-Pooling，簡稱：MP）。
1994 年	早期競賽獲勝的神經網絡。
1995 年	長短期記憶（Long Short-Term Memory，簡稱：LSTM）。
2003 年	更多獲獎、創紀錄的神經網絡與成功的深度神經網絡。
2006~2007 年	深度信念網絡（Deep Belief Network，簡稱：DBN）、針對 MPCNN 與 LSTM 提升 CNN 的 BP 與 GPU。
2009 年	由 RNN 和 MPCNN 贏得首次正式比賽。
2010 年	GPU 上的反向傳輸打破了 MNIST（Modified National Institute of Standards and Technology，簡稱：MNIST）的記錄。
2011 年	GPU 上的 MPCNN 實現視覺性能。
2012 年	用 Hessian Free 方法優化 RNN。
2013~	更多比賽和記錄。

藏層中含有許多節點（Nodes）或是稱為神經元（Neurons），節點的數量可由設計者所控制，通常是透過嘗試錯誤法來決定節點的數量。隱藏層的層數也是隨著分類問題而有所不同，隱藏層層數越多通常可以處理越繁雜的問題，但是相對的也會占用

較多電腦的運算資源，讓訓練過程時間變得冗長。輸出層的功能為將最後的分類結果傳遞出去。

## （二）卷積神經網路

卷積神經網路（Convolutional Neural Network, CNN）是一種適合用於處理圖

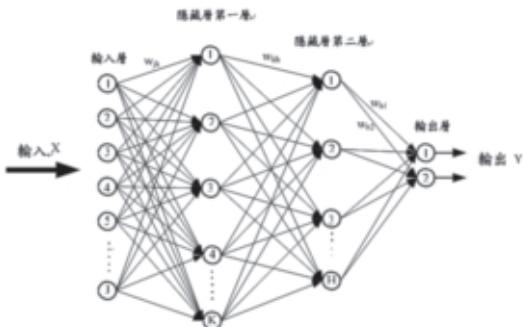


圖 3-2 多階層感知器基本架構

資料來源：



像、影像辨識問題的神經網路架構，由許多不同功能的隱藏層（Hidden Layer）組成，包含卷積層（Convolutional Layer）、池化（Pooling）、全連接層（Fully Connected Layer）。在實際運用演算中多將一維訊號

當作二維圖片，利用卷積神經網路的特性擷取特徵進行學習。

卷積層（Convolutional Layer）的主要作用為特徵擷取，可學習到圖像的局部特徵。擷取特徵的原理是使用稱作卷積核（kernel map）的移動窗格去對原圖進行濾波，也就是將卷積核與原輸入矩陣內的局部區域進行內積，如圖 3-3，卷積核每次移動的距離稱為步長（stride），不同的卷積核會有不同的參數，目的是為了擷取輸入圖片中不同的特徵。原始圖像每經過一個卷積核濾波後，會產生一張濾波後的特徵圖，因此卷積核的個數就是卷積層的厚度，如圖 3-4 所示。

有時會希望卷積後的輸出圖片能維持大小，此時會將原輸入圖補零（zero

1	1	0	1	1	0	0
0	1	1	0	0	1	1
0	1	0	0	1	0	1
1	0	1	1	0	1	1
1	1	0	0	0	0	1
0	1	0	1	1	0	0
1	0	1	0	1	1	1



0	0	1
0	1	0
1	0	0



1				

1	1	0	1	1	0	0
0	1	1	0	0	1	1
0	1	0	0	1	0	1
1	0	1	1	0	1	1
1	1	0	0	0	0	1
0	1	0	1	1	0	0
1	0	1	0	1	1	1



0	0	1
0	1	0
1	0	0



1	3	1	0	2
3	0	1	3	1
1	2	2	0	2
1	2	0	2	2
2	0	2	1	2

圖 3-3 圖像經卷積核濾波示意圖

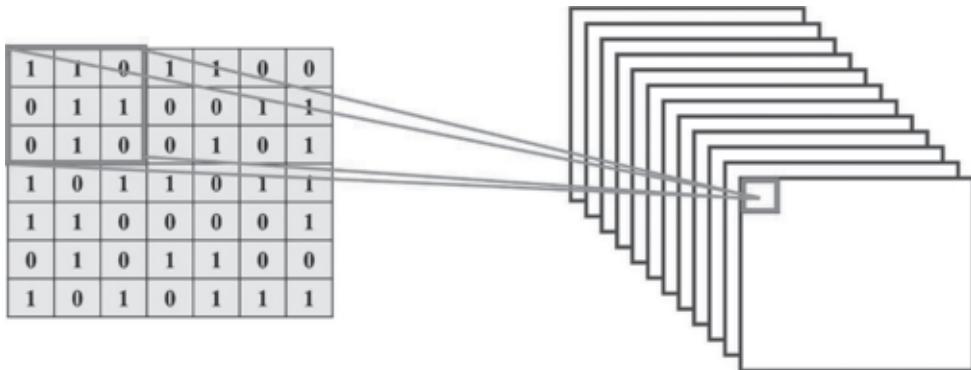


圖 3-4 圖像經多個卷積核濾波示意圖

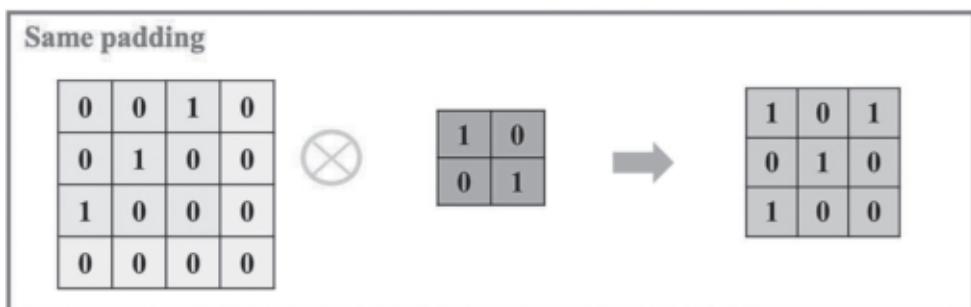
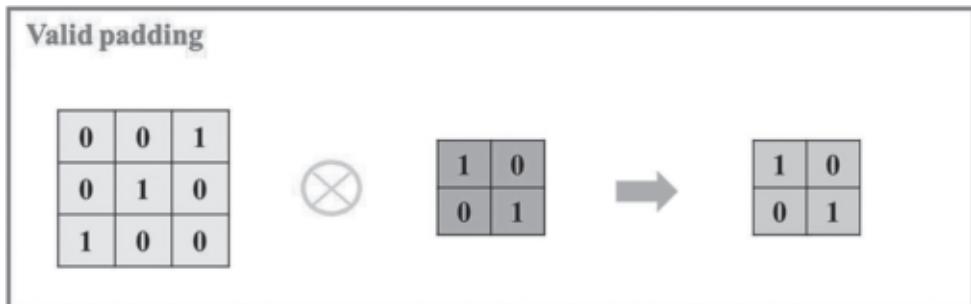


圖 3-5 Valid padding 和 Same padding 示意圖

padding)，經過補零控制輸出圖片矩陣大小不改變的方法稱作 same padding；反之若輸出圖片矩陣改變大小稱為 valid padding，如圖 3-5。

池化的主要作用為將輸入圖像降維，

留下較能夠幫助分類的特徵，依照保留特徵的方法分為平均池化（Average Pooling）及最大池化（Max Pooling）。池化之作用原理為取出原圖片矩陣中各個區域中的最大值或是平均值，若是取最大值則稱最大

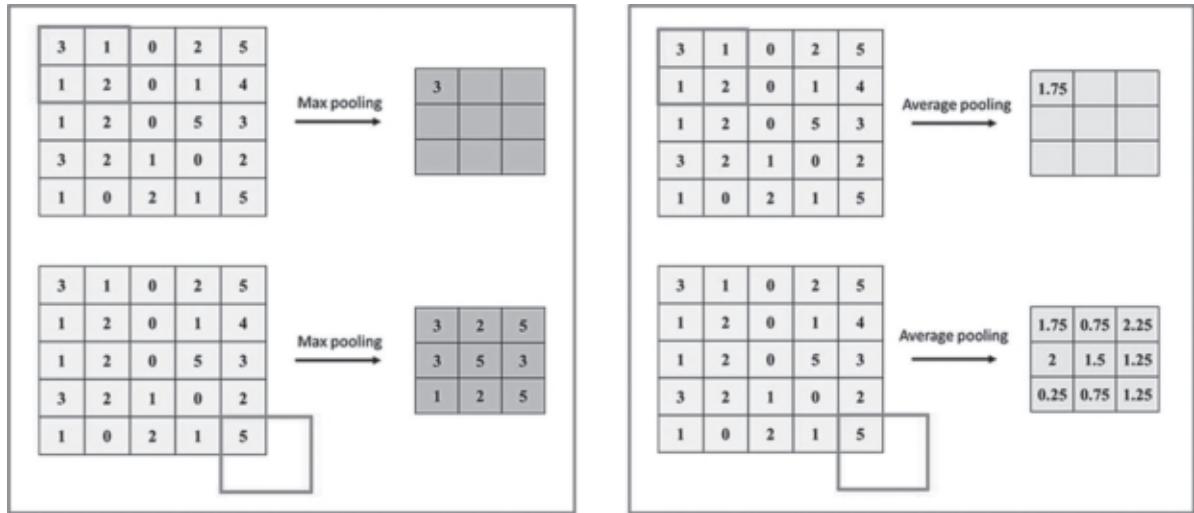


圖 3-6 最大池化與平均池化示意圖

池化；若取平均值稱為平均池化。以下圖 3-6 為例，使用一個二乘二的濾波器去對一個五乘五的矩陣池化，步長（stride）為（2,2），則最大池化會取出濾波器範圍中數值最大的特徵，如圖 3-6 左側圖所示；而平均池化則是會取濾波器範圍中的平均值，如圖 3-6 右側圖所示。

全連接層（fully-connected layers）就是傳統的神經網路，將前面經過一系列卷積層和池化層提取出的局部特徵統整起來，進行最後的分類，三層以上的全連接層即為上節（一）所說明的多層感知器（Multilayer Perceptron, MLP）。圖 3-7 為網路實際利用卷基層演算後圖形顯示的圖例。

### （三）循環神經網路與長短期記憶網路

現今最常用來使用處理時間序列問題的神經網路之一是循環神經網路（Recurrent

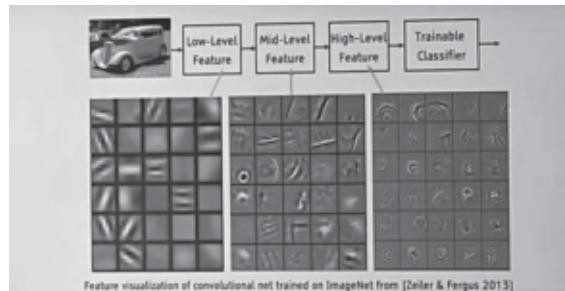


圖 3-7 成為參數被網絡學習後的圖形顯示

資料來源：

原文網址：<https://kknews.cc/tech/3xok93.html>

Neural Networks, RNN），而目前深度學習相關研究人員所使用的循環神經網路一般都是指長短期記憶網路（Long Short-Term Memory Network, LSTM）。長短期記憶網路是循環神經網路的一種，一個基本的長短期記憶網路結構中包含了輸入閥（Input

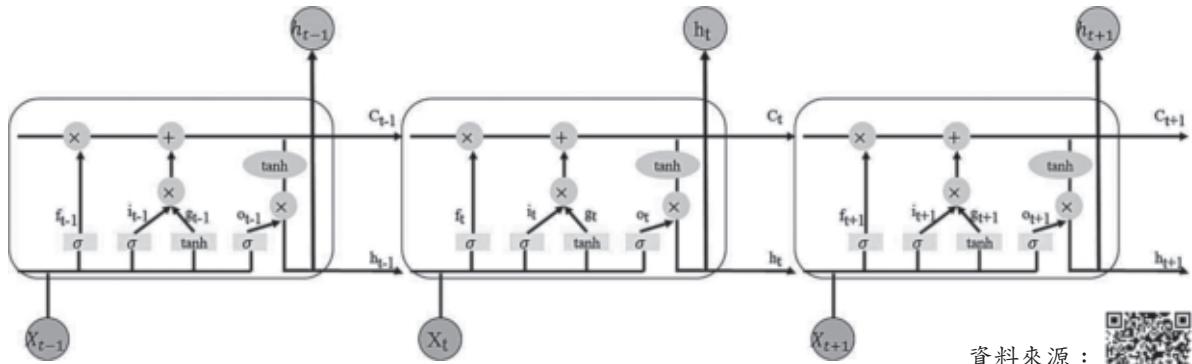


圖 3-8 長短期記憶網路基本結構圖

資料來源：



Gate)、輸出閥 (Output Gate)、遺忘閥 (Forget Gate)，結構如下圖 3-8 所示。輸入閥是用來控制輸入是否要被寫入記憶 (Memory) 裡面；輸出閥是用於控制是否可以把輸出值讀出來；遺忘閥則是用來控制什麼時候要把之前儲存的記憶刪掉。輸入閥、輸出閥、遺忘閥的作用原理是將輸入以及前一次演算的輸出值乘上一組權重，分別通過激勵函數（通常使用 Sigmoid 函數）來調控讓輸入通過的程度。透過三個閥的調控，可以彈性控制在不同時間，神經網路對時間序列資料的記憶效果。

#### (四) Inception 網路架構

隨著深度學習日益受到重視，神經網路發展迅速，新架構不斷推陳出新，其中伴隨來的是龐大的參數量，造成演算緩慢、演算時占用電腦硬體大部分運算資源，而 Google 的開發團隊提出的 GoogLeNet 解決了此問題，其中扮演重要角色的就是 Inception 網路結構。

Inception 網路結構使用了  $1 \times 1$  卷積層 (Convolution)。其具備了可以降維，減少參數量的特性，幫助神經網路在架構加深的同時能控制參數量。舉例來說：若前一層輸出矩陣大小為  $20 \times 20 \times 128$ ，要經過 256 個 kernel 的  $5 \times 5$  卷積層，其步伐 (stride) 為 1、填充 (padding) 為 same padding，因此輸出數據為  $20 \times 20 \times 256$ ，其中用到之卷積層參數量為  $128 \times 5 \times 5 \times 256$  個，如下圖 3-9 所示。若在中間加上一層 32 個 kernel 的  $1 \times 1$  卷積層，則輸出數據同樣為  $20 \times 20 \times 256$ ，但是卷積層參數量減少為  $128 \times 1 \times 1 \times 32 + 32 \times 5 \times 5 \times 256$  個，減少了將近 4 倍之多，則演算速度勢必提高。

除了使用  $1 \times 1$  卷積層，Inception 網路結構還將神經網路結構加寬，將多種不同大小的卷積層和最大池化堆疊在一起，如圖 3-10，增加神經網路對前一層輸出在各種尺度的特徵萃取能力，相當於提取腦電訊號在不同頻帶的特徵。

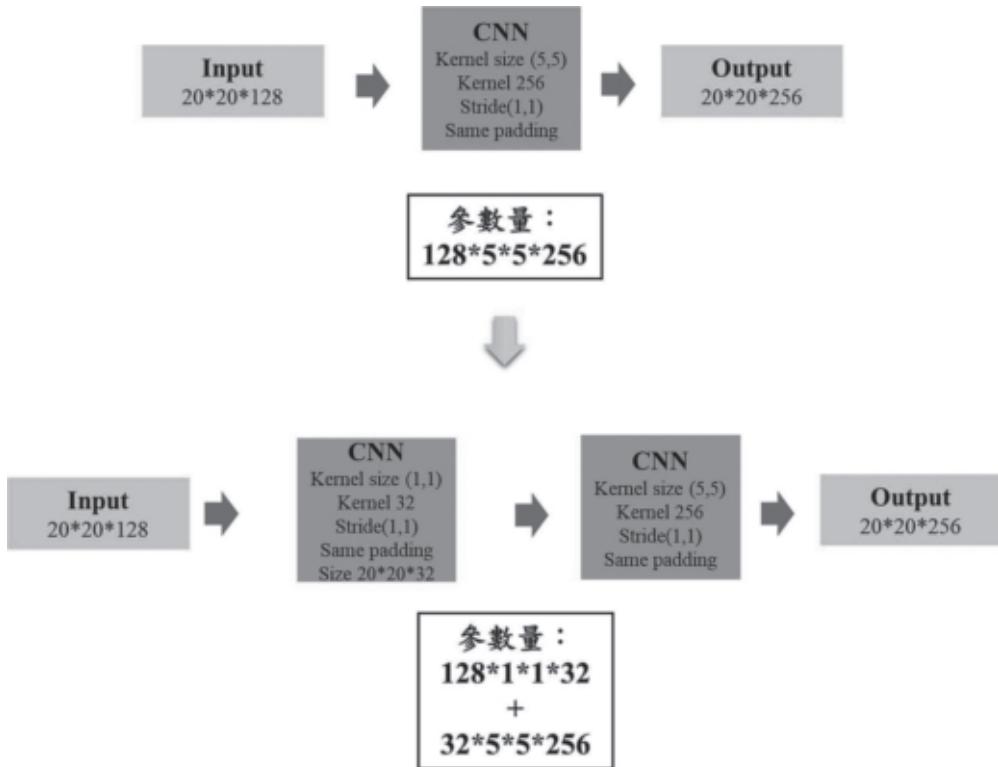


圖 3-9 使用  $1 \times 1$  卷積層（Convolution）示意圖

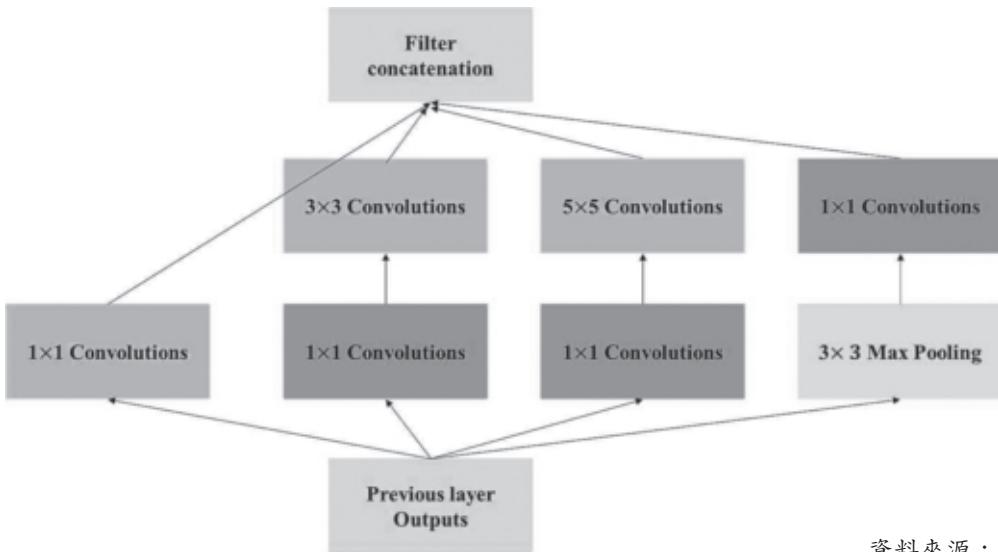


圖 3-10 Inception 基本結構

資料來源：



### 三、深度學習的應用

深度學習的應用領域相當廣泛，包括：圖像識別、語音識別、自然語言處理、目標檢測、藥物發現、基因學以及廣告點擊率預估等。深度學習為現代社會的許多方面提供了動力，從網路檢索到社交網絡上的內容過濾，再到電子商務網站的推薦，以及相機和智慧型手機等產品的功能，都能發現深度學習的蹤影（LeCun, Bengio, & Hinton, 2015）。以下針對深度學習的應用進行舉例與說明。

#### （一）圖像識別

圖像識別的發展已有相當長的一段時間，圖像識別也是深度學習最早嘗試的領域，它可以應用的範圍相當廣泛，像是基於內容的圖像檢索、最近相當熱門的臉部辨識以及近期發展相當快速的醫學圖像分析等。

在臉部辨識方面，由於複雜的臉部變化，預測臉部屬性是具有挑戰性的，Liu 等人提出了一個新的屬性預測的深度學習框架，該框架包含了 CNN、LNet 以及 ANet，這個框架不僅比現有技術大得多，而且還揭示了有關人臉表徵的價值，這將有益於未來的人臉辨識和深度學習研究（Liu, Luo, Wang, & Tang, 2015）。Parkhi 等人的目標是從單張照片或從視頻中追蹤一組人臉，最近在這方面取得的進展是由於兩個因素：使用卷積神經網絡端到端的學習與超大規

模訓練數據集的可用性，Parkhi 等人展示了一個大規模的數據集（2.6M 圖像，超過 2.6 萬人）並透過深度網絡訓練和人臉識別的複雜性來介紹方法和程序（Parkhi, Vedaldi, & Zisserman, 2015）。Wang 等人則是將臉部辨識用於安全監控，由於安全監控實際應用的需求，在現實世界的監控影像中進行穩健性的人臉識別是一個具有挑戰性的重要問題，儘管當前的臉部辨識系統在相對受限的場景中表現良好，但它容易因為姿勢、照明或臉部表情的變化而有所影響，Wang 等人因而提出了一種透過深度學習在實際監控影像中進行臉部辨識的方法（Wang, Bao, Ding, & Zhu, 2017）。臉部辨識除了許多相關的研究外，也有一些實際的應用，像是 iPhone 的 Face ID、Google 相簿中的人臉辨識功能以及人臉辨識門禁系統等。

在醫學圖像分析方面，近年來也有相當多的研究，深度學習演算法，特別是卷積神經網絡，已經迅速成為分析醫學圖像的首選方法。Litjens 等人回顧了與醫學圖像分析有關的主要深度學習概念，他們調查了使用深度學習進行的圖像分類、目標與病變檢測以及器官與子結構分割等任務，並對每個應用領域的研究提供簡明的概述，包括：神經、大腦、視網膜、胸部、乳房、心臟、腹部以及肌肉骨骼等的圖像分析。從他們回顧的 308 篇論文中可以看

出，深度學習已經應用到醫學圖像分析的各個面向，而且發展相當迅速，絕大多數的貢獻都是在 2016 年左右發布的（Litjens et al., 2017）。Gulshan 等人就是以深度學習進行的圖像分類來檢測糖尿病視網膜病變，它以高靈敏度和高特異性鑑別糖尿病視網膜病變或雙眼黃斑水腫，這種檢測糖尿病視網膜病變的自動化系統具有多種優勢，包括解讀的一致性、高靈敏度、特異性以及即時的結果報告。但這類型的研究依然有它的限制，Gulshan 等人所使用的參考標準是所有眼科醫生分級的主要決定因素，這意味著大多數眼科醫師無法識別的細微結果的圖像可能無法完成，另一個限制來自深度神經網絡的性質，神經網絡僅具有圖像和相關等級，沒有明確定義特徵，最後，尚未解決的問題是，眼科醫師使用的界面設計和分級設置是否會影響臨床表現，這還需要進一步的實驗來解決（Gulshan et al., 2016）。近期，醫學領域應用圖像識別的研究相當多且發展相當迅速，但是，應用於醫學的圖像識別往往需要進一步的評估和驗證，才能明確瞭解實驗研究與實際臨床狀況的差距。

深度學習可以用來識別照片中的文字。一旦識別了，文字就會被轉成文本，並且被翻譯，然後圖片就會根據翻譯的文本重新創建。這就是我們通常所說的即時視覺翻譯。假設你現在身處一個非母語國

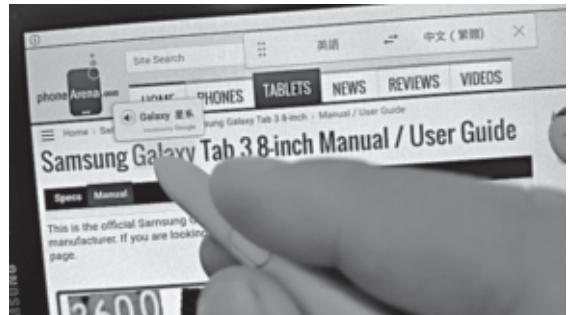


圖 3-11 深度學習在即時視覺翻譯的實際應用圖示

資料來源：



家，你也不用擔心，只要下載一個谷歌翻譯這樣的 app，你就可以大膽往前沖。利用這樣的 app，你可以去翻譯路標或門店名。這類軟體之所能實現這些目標，其實就是因為深度學習的應用（如圖 3-11 所示）。

## （二）語音識別（圖 3-12）

語音識別也是深度學習領域中發展相當久的技術，Facebook、Google、IBM、微軟以及百度等企業針對語音識別技術的應用與發展投入了不少心力，深度學習正成為工業領域語音識別的主流技術。以微軟為例，Deng 等人概述了自 2009 年以來微軟語音研究人員在該領域的工作，並介紹了最近的進展，為當前的深度學習技術的基本功能和局限性提供了解決方案，他們提供了最近將深度學習方法用於提升語音技術和相關應用的實驗樣本，與其他近年來的工作表明，深度學習是一項強大的技

## 語音辨識系統的基本架構

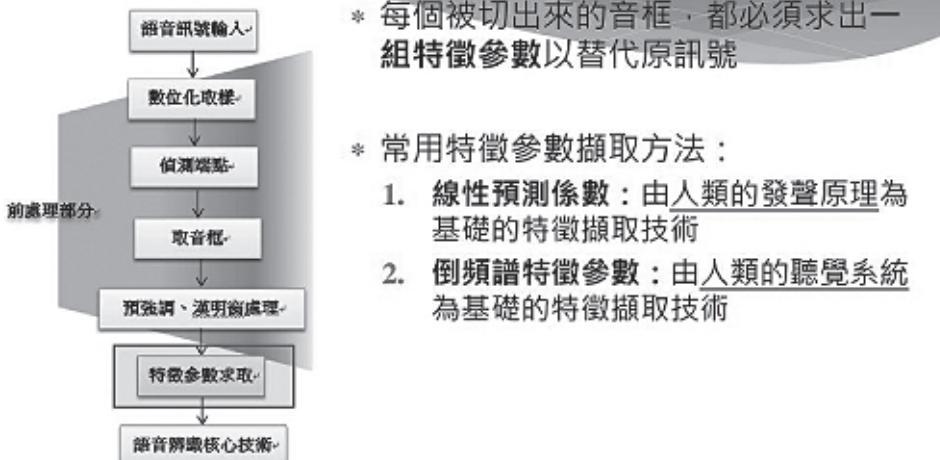


圖 3-12 語音辨識基本架構圖

資料來源：



術，他們今後在深度學習研究方面的工作主要針對三個基本方向：更有效的深層架構和學習算法、使用越來越大的數據集對深度學習模型進行訓練以及將深度學習模型的應用擴展到其他語言和語言處理領域以及其他領域（Deng et al., 2013）。

### (三) 無人機

深度學習最近在各種機器人任務方面表現相當出色，其在實際環境中獲取複雜數據，並有出色的學習能力，這讓它非常適合多種自主機器人的應用。無人機目前正廣泛應用於安全、監控、災難救援、包裹運輸以及倉庫管理等多種民用任務，Carrio 等人評估無人機深度學習的應用，包

括開發與其性能和局限性，並描述無人機深度學習應用的主要挑戰。現在的機器人能夠執行複雜的操作，但是並不容易適應環境中的變化，在這種情況下，就必須依靠經驗學習，而深度學習就是解決這個困境相當有用的技術（Carrio, Sampedro, Rodriguez-Ramos, & Campoy, 2017）。

實際應用以大量使用基於深度學習的對象檢測方法來自動分析航空影像，以用於農業，建築，公共安全和安全等領域。台灣科技部「2018 未來科技展」( Future Tech Expo, FUTEX 2018 ) 已於 2018 年 12 月 13 至 15 日於臺北世貿三館舉行。展覽有三大主題，包括量子電腦、AI 人工智能



圖 3-13 無人機拍攝畫面可即使辨識和計算數量圖示

資料來源：

與電子光電應用、以及精準醫療應用等前瞻性技術展示。當中亦不乏無人載具技術應用展出，如台灣大學資訊工程學系及現場展出「大規模無人機物件偵測卷積網路技術」，這個技術可以透過無人機高空拍攝，計算地面物件數量。示範影片展示了拍攝停車場，可以即時計算現場有多少車輛，可藉此回報尚有多少空位，比起人力計算來得快捷和有效率。這個技術透過深度學習辨識車輛，無論靜止、移動中或陰影下的車輛都可辨識。據悉除了辨識物件的形狀大小，也加入了特定排定形狀的數據，而且這技術的錯誤比起世界上其他頂尖的演算法還要來得低。這技術目前尚未命名，未來除了可應用於航空拍攝計算，如畜牧業計算牲口數量、人群計算等，亦

可應用在醫學，如透過顯微鏡拍攝計算紅血球等細胞數量。

#### (四) 金融市場預測

大數據分析和深度學習是數據科學的兩大重點，大數據對於需要收集大量數據的組織非常重要，因為許多組織都在收集大量特定領域的資訊，這些資訊可能包含國家情報、網路安全、詐欺檢測、市場營銷以及醫療資訊等有用的資訊，而深度學習幫助了大數據的收集、分類以及應用，這使得深度學習成為極具價值的工具 (Najafabadi et al., 2015)。

舉例來說，人類沒有辦法預測未來某個時期可能表現良好的股票，但是，深度學習對於這樣的問題可能是有用的，深度學習提供了使用大型數據集來優化預測性

能的框架，因此，深度學習框架非常適合金融領域的實踐與理論，Heaton 等人使用深度學習分層模型來解決財務預測和分類問題，期望能透過深度學習的方法來提高傳統應用的預測性能，甚至是能夠顯著提升（Heaton, Polson, & Witte, 2017）。

有鑑於此，基於大數據分析和深度學習技術，目前金融產業致力推廣「機器人理財」的方法，提供客戶自動化、客製化投資的網路平台。利用使用者提供的資料，衡量規劃合適的策略及風險承受度，並利用演算法來作分散式資產配置，並定期追蹤市場狀況，自動化再平衡投資組合。簡單來說，機器人理財排除了傳統的人為主觀影響，完全以電腦演算法管理配置資產，保持絕對的理性與紀律，完美運用投資理論於實際金融市場操作（如圖

3-14）。

### （五）情感分析

社群媒體使用者生成內容的情感分析對於許多社群媒體分析任務來說相當重要，研究人員依靠文本情緒分析來開發系統，以預測政治選舉與衡量經濟指標等。最近，社群媒體使用者越來越多地使用圖像和影像來表達他們的觀點與分享他們的經驗，這種大規模視覺內容的情感分析有助於更好地提取使用者對於事件或主題的觀點，以便從視覺內容中預測情緒。

You 等人為圖像情感分析設計卷積神經網絡架構，透過使用情感分析方法標記 Flickr 圖像來獲得 50 萬個訓練樣本，此外，他們透過使用少量手動標記的 Twitter 圖像來提高 Twitter 圖像的性能，結果表明，提出的卷積神經網絡在圖像情感分析方面可

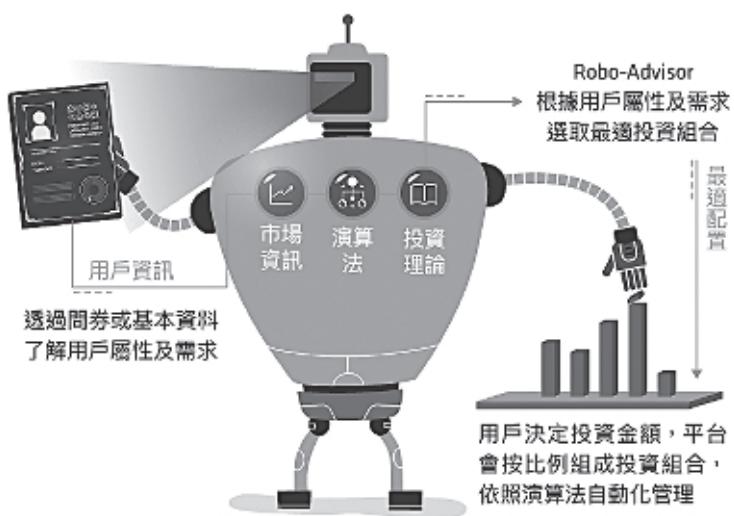


圖 3-14（圖片來源：StockFeel 股感知知識庫）

資料來源：





以取得更好的性能（You, Luo, Jin, & Yang, 2015）。Sohangir 等人則針對股票市場進行情緒分析的研究，股票市場是增加財富的熱門場所，但何時買入與賣出股票，或購買哪些股票的基本問題尚未解決，透過如 StockTwits 與 SeekingAlpha 等金融社交網絡的普及，世界各地的投資者都有機會收集和分享他們的經驗，Sohangir 等人調整深度學習模型以改善 StockTwits 的情緒分析性能，他們應用了幾種神經網絡模型來發布 StockTwits 的股票市場意見，研究結果表明，深度學習模型可以有效用於金融情緒分析，而卷積神經網絡是預測 StockTwits 數據集中使用者情緒的最佳模型（Sohangir, Wang, Pomeranets, & Khoshgoftaar, 2018）。

總結上述，深度學習的發展已經有一段時間了，深度學習的發展算是相當出色，應用的領域也相當多元與廣泛，對實際應用的貢獻來說，深度學習可能是機器學習領域最近這十年來最成功的研究方向，但是，深度學習的發展仍存在著一些問題，像是理論、建模以及工程上的問題，尤其是深度學習的理論研究，這是相當困難的，舉例來說，它缺乏關於深度神經網絡中目標函數幾何的知識，或是難以了解為什麼某些架構比其他架構更好，所以，即便現在的深度學習技術相當出色，但它仍然有很大的空間值得去研究與發展

（余凱等人，2013；Carrio et al., 2017）。

## 肆、結論與未來展望

2018 年 1 月才正式開幕的亞馬遜無人超市 Amazon Go，標榜「商品拿了就走」的快速消費體驗，一開張就在全球引起轟動，不僅吸引大批來自世界各地的滿滿人潮，帶著好奇心，等著進店消費，它所掀起的無人零售的新型態，更成為科技圈最火紅的話題，甚至在全球掀起一股無人商店的風潮，包括中國、日本、韓國與臺灣等連鎖業者皆相繼投入，Amazon Go 強調沒有排隊人潮，也不用結帳櫃臺，消費者只要具備 Amazon 帳號、下載 Amazon Go App，以手機 QR 條碼刷進閘門即可開始消費。整間店僅在入口處設置 6 個自動閘門管控人員進出，其中 4 道門只進不出，另外 2 處則當作出口通道，離開時皆從此通過（如圖 4-1）。



圖 4-1 手機成為進 Amazon Go 消費的唯一憑證

這間無人商店就設在西雅圖市中心，距離亞馬遜總部辦公室僅僅幾步之遙，員工從樓上辦公室走到 1 樓就看得到。店內約 50 坪大，室內空間和臺灣的便利超商差不多。亞馬遜一共耗時 4 年多，才完成建置了這間無人超市，並在 2016 年底首度對外亮相。不過，一開始僅對內部員工開放，並沒有對外營運，在歷經長達 1 年內部測試，Amazon Go 才於 2018 年 1 月 22 日正式啓用營運。現在即使是一般消費民眾，也能進到店內親自體驗。該店目前開放時間為早上 7 點到晚上 9 點，且僅平日營業，假日休息不營業。走進這家無人商店，室內不僅空間明亮，還有寬敞的走道與挑高設計，Amazon Go 強調沒有排隊人潮，也不用設立結帳櫃臺，僅在入口處設置 6 個自動閘門管控人員出入，其中 4 道門只進不出，另外 2 處則當作出口通道，離開時皆從此通過。在閘機上面還嵌入一個條碼掃描機，民眾以手機打開 Amazon Go App 對著閘門輕掃 QR 條碼就可進入，離開時站在出口閘門前，門就會自動開啟，不需要刷門就可拿著商品自動離去。顧客離店時只須將商品拿在手上，或是放進店家免費提供的橙色購物袋，或牛皮紙袋，就能從容離去，不需要排隊結帳。該商店會根據顧客離店時所攜帶的物品品項和數量自動計費，再從該帳戶事先綁定的信用卡扣款。

從顧客進入商店拿取商品到離開結帳的整個過程，皆結合電腦視覺辨識技術，並自動追蹤顧客取貨消費行為，Amazon Go 無人商店的核心技術就在店內結合了電腦視覺辨識技術，並搭配多種環境感測器（Sensor Fusion）等運用於自動駕駛的技術，打造出該公司稱為 Just Walk Out（拿了就走）的無人零售新模式，它運用店內的相機、感測器，來追蹤消費者拿了或放回哪些商品，並根據拿走的商品自動結帳，徹底實現全自動化的無人自助結帳。亞馬遜形容，這就像是把線上購物模式整套搬到了實體商店，顧客邊走邊選購商品時，身旁會有個看不見的虛擬購物車，自動跟著一起移動，顧客每從貨架上拿取一件商品，等同於線上將該商品計入該用戶帳戶的虛擬購物車裡，而將看完商品放回架上的動作，也等於是直接取消訂購，這時，商品就會從購物車裡被移除。一整個取貨購買的過程，看起來就像在線上購物，就連後續收到帳單的方式，也都頗為類似，例如都會收到電子帳目明細，只是這裡是以手機做為媒介，顧客拿著商品離開後，手機上就會收到帳單通知，從手機即可查閱購物明細，甚至還可查看造訪停留時間。兩者最大不同在於，顧客在這裡發生的所有購買行為都是線下進行，可以摸得到實體商品，而且拿了就能走，店家會自動完成結帳，寄發帳單。另一個不同之

處，這裡不只消費者自己一人能夠進店購物，甚至也能允許帶自己的家人、朋友進店消費。只須以手機刷門先讓朋友或家人依序進入後，最後用戶自己再刷門進入就行，如此一來，這些家庭成員或友人在店內的一舉一動，也都會被自動追蹤記錄，而這些人拿走的商品，最後會記在該用戶的帳單上，而不會與其他人搞混。

Amazon Go 能完成無人自助結帳服務的關鍵，主要是它利用超過 50 個以上的深度感測相機組成的相機陣列，打造出一套結合電腦視覺的影像辨識系統，能自動偵測商品、分析消費者動作，以及追蹤在店內移動的路徑，以判斷顧客拿取了何種商品。這些相機被懸掛在天花板下方，並以各種不同角度，對準下方的走道或是貨架上的商品，以偵測追蹤消費者的一舉一動。這種靠影像辨識偵測商品、顧客舉止動作，自動判斷拿了何種商品的技術應用，Amazon Go 至少以超過 50 臺相機組成的相機陣列，打造出這套影像辨識系統，這些相機被裝在一個個黑盒子內，懸掛在有著裸露鋼架和管線的天花板正下方，只要一抬頭就能看到。天花板上的相機形狀也有所不同，有的造型是正方形，也有長方形的造型，如果仔細近看，可以發現其中長柱體的相機配有上下兩顆鏡頭，推測應具備 3D 深度感測的能力，天花板上不少都是屬於這類的相機。另外，相機放置的

角度也有些許不同，但共同之處，就是都將鏡頭對準下方的走道或是貨架上的商品，以偵測追蹤消費者的一舉一動（如圖 4-2）。這套影像辨識系統還結合深度學習技術，不僅可以正確的辨識現在站在貨架前的這個人是誰，拿了什麼商品，就算同時有數個人並排也都能夠辨別，例如，可以知道誰剛剛拿走了優格或蛋糕，或是誰在查看商品，甚至還可進一步依據顧客的動作變化，來預測對方接下來的購物行



圖 4-2 賣場內超過 50 臺深度相機陣列是  
偵測追蹤的關鍵

為，比如，準備拿起商品，還是正要將其放回原處，藉此來加快系統處理的反應時間。

店內的貨架上也都整合多功能感測器作為輔助。據了解，Amazon Go 在每個貨架上，皆設有重量感測器，能偵測架上商品的重量，並比對影像辨識的結果，以提高辨識準確度。另外，透過分析這些感測器所蒐集的數據，也能用來提高人員作業效率，像是當商品即將銷售一空時，就會自動發送訊息通知店員補貨，亦或是商品放錯位置或被擺在不對的貨架上時，也會發出警報，店員從手持裝置上就能查看，將商品重新放到正確的貨架上。不過商品貨架並未導入 RFID 電子標籤，仍需要人工更換標籤，這也是 Amazon Go 日後可以加以改進之處。Amazon Go 在辨別商品、分析顧客行為動作的出色能力，甚至還能夠用來預防它免於遭竊，即使是消費者想要試圖欺騙，也很難躲過它的法眼，甚至比店員都還聰明，而且無死角，就算沒有店員在旁隨時緊迫盯人，憑著店內大量攝影機和感測器，就能追蹤顧客從貨架上拿下或放回哪些商品，Amazon Go 利用結合電腦視覺的相機，搭配貨架上的多種環境感測器，來偵測與追蹤消費者拿了或放回哪些商品。顧客每從貨架上拿取一件商品，就會將該件商品計入該用戶帳戶的虛擬購物車裡，若是將商品放回架上，則會自動



圖 4-3 電腦視覺追蹤消費者取貨購買行為

將它自購物車裡移除。就算拿完商品要放回時不小心擺錯了位置，並不是原先商品放置的位置，也一樣能夠正確辨別，不會將該商品列入計費（如圖 4-3）。

顧客拿著商品離開後，經過 5 到 10 分鐘，手機上就會收到帳單通知，消費者從手機上可查看商品購物明細以及到店停留時間，若是帳單上記載品項與實際購買的商品不符，直接從手機上的帳目明細，就可將商品移除，即時完成退款（如圖 4-4）。

在影像辨識上還結合深度學習，不僅



圖 4-4 顧客離店後自動收到帳目明細

可以分辨貨架前的人群，現在拿貨或是移動的路徑，甚至還可進一步依據顧客的動作變化，來預測對方接下來的動作，會是拿商品還是將其放回原處（如圖 4-5）。靠著這套影像辨識系統也能識別不同商品，迄今在 Amazon Go 店內可供辨識的販售商品超過 1,600 項，包括零食、冷飲、熟食、酒類飲品，以及雜貨等都能正確辨別。比較特別的是，為提高商品的辨識率，店內對於採用相同外盒包裝容易辨識不易的鮮食產品，還會加上專用識別設計，讓天花板上的相機更容易辨識顧客拿了哪種食品（如圖 4-6），並可確保鮮食產品的新鮮度。

在 2019 年 8 月，日本電信公司 NTT 旗下的子公司 NTT DATA 與中國雲拿科技（cloud pick）合作，在東京開設一家無人實驗商店，推出「無收銀台」的結帳模式。改善以往無人商店需用額外的支付系統，在此商店購買商品不須掃描任何條碼，離開商店後就會直接自動結帳，消費者只需下載特定應用程式就可以自動支付款項。不僅如此，NTT DATA 的最終目標是促進無人商店在日本的實際應用，預計在 2020 年前引進 1,000 家便利店、藥店、超市等。無人商店早在之前就有持續的發展，像是美國亞馬遜的「Amazon GO」、中國淘寶的「淘咖啡」及我國的「7-11 X-STORE」所面臨的問題一樣，無人商店成為零售業者的解決方案，從買賣商品到結帳都不須花費



圖 4-5 商店可靠深度學習技術預測店內人潮動向



圖 4-6 鮮食品附特殊標記能提高辨識率

資料來源：(掃描 QR Code 連結觀看說明影片)



人力，最多只需對產品進行補貨，如此一來就可以有效緩和人力不足的問題，然而，缺點是為維持無人商店的營運，店內設備加上店面維護所費不貲，無人商店業者必須從技術層面下手，才能有效降低成本，而且儘管無人商店帶給人們便利，但也磨滅了店員及顧客間熱情的招呼，亞洲地區所重視「集體文化」，對於人與人之間的互動也影響亞洲地區服務業的發展，「顧客至上」一直是亞洲服務業的至高準則，各業者巧思盡出，只為能提供顧客更貼心與安心的服務，讓消費者留下美好消費回憶，但這也是無人商店所欠缺的「溫度」。

科學家指出，人工智慧在未來 10 年內即將就會改變你我的生活方式，有鑑於無人商店帶給商家銷售和統計的便利，卻也讓消費者不再有與人面對面門市零售的體驗，甚而人工智慧對產業的衝擊和工作職務的取代等等，人工智慧帶給我們生活的便利和未來產業結構的改變，期許藉由本文拋磚引玉，未來能有更多相關的研究和探討，讓印刷產業也能與人工智慧科技與時俱進。

## 參考資料

1. Anand Rao, Joseph Voyles, & Pia Ramchandani (2017). Top 10 artificial intelligence (AI) technology trends for 2018. 上網日期 2019 年 6 月 20 日，檢自：<http://usblogs.pwc.com/emerging-technology/top-10-ai-tech-trends-for-2018>
2. Aslani, A., Mazzuca-Sobczuk, T., Eivazi, S., & Bekhrad, K. (2018). Analysis of bioenergy technologies development based on life cycle and adaptation trends. Renewable Energy, 127, 1076-1086.
3. Carrio, A., Sampedro, C., Rodriguez-Ramos, A., & Campoy, P. (2017). A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles. Journal of Sensors, 2017.
4. Chen, C. C., Yang, K. H., Kao, H. Y., & Ho, J. M. (2008, March). BibPro: A Citation parser based on sequence alignment techniques. In Advanced Information Networking and Applications-Workshops, 2008. AINAW 2008. 22nd International Conference on (pp. 1175-1180). IEEE.
5. Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. Journal of the American Society for Information Science and Technology, 62(7), 1382-1402.
6. Espacenet. (2018). Espacenet patent search. 上網日期 2019 年 8 月 4 日，檢自：<https://www.epo.org/searching-for-patents/technical/>

- espacenet.html#tab-1
7. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ...& Kim, R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22), 2402-2410.
  8. Kim, D. H., Lee, B. K., & Sohn, S. Y. (2016). Quantifying technology-industry spillover effects based on patent citation network analysis of unmanned aerial vehicle (UAV). *Technological Forecasting and Social Change*, 105, 140-157.
  9. Kim, J., Lee, J., Kim, G., Park, S., & Jang, D. (2016). A hybrid method of analyzing patents for sustainable technology management in humanoid robot industry. *Sustainability*, 8(5), 474.
  10. Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3730-3738).
  11. Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, 2(1), 1.
  12. Sohangir, S., Wang, D., Pomeranets, A., &
  - Khoshgoftaar, T. M. (2018). Big Data: deep learning for financial sentiment analysis. *Journal of Big Data*, 5(1), 3.
  13. Tkaczyk, D., Collins, A., Sheridan, P., & Beel, J. (2018, May). Machine Learning vs. Rules and Out-of-the-Box vs. Retrained: An Evaluation of Open-Source Bibliographic Reference and Citation Parsers. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (pp. 99-108). ACM.
  14. Amazon Web Services (AWS). (2019). What is Artificial Intelligence? 上網日期 2019 年 8 月 26 日，檢自：[https://aws.amazon.com/machine-learning/what-is-ai/?nc1=h\\_ls](https://aws.amazon.com/machine-learning/what-is-ai/?nc1=h_ls)
  15. Business Next. (2019). Artificial Intelligence (AI) . 上網日期 2019 年 9 月 8 日，檢自：<https://www.bnnext.com.tw/search/tag/AI>
  16. VantagePoint. (2018). Products. 上網日期 2019 年 7 月 4 日，檢自：<https://www.thevantagepoint.com/products/4-products/vantagepoint/15-turn-information-into-knowledge.html>
  17. From Wikipedia. (2019) Deep learning. 上網日期 2019 年 8 月 24 日，檢自：[https://en.wikipedia.org/wiki/Deep\\_learning](https://en.wikipedia.org/wiki/Deep_learning)

中央印製廠 / 劉毓容