

FOREWORD

推薦序

蔡宗翰教授與裴有恆老師致力於 AI 與 AIoT 的教學研究及服務，這本書《AI+AIoT 概論：寫給大學生看的 AI 通識學習》乃探討人工智慧的概念、類型、技術和應用，是一本相當棒的 AI 入門書籍。

人工智慧是一門研究如何使機器具有智慧的學科，目標是使機器具有像人類一樣的智慧。人工智慧可分為弱人工智慧、強人工智慧和專家系統三種類型。人工智慧的技術有機器學習、自然語言處理、電腦視覺等。人工智慧的應用廣泛，包括自動駕駛、智慧家居、語音辨識、語音合成、醫療、客服等。本書將透過舉例來解釋人工智慧的各種技術和應用，並探討人工智慧的局限性和潛在挑戰。同時，本書也將討論人工智慧的道德和倫理問題，包括人工智慧決策的透明度和可追溯性、人工智慧的偏見和歧視、人工智慧對就業市場的影響等。

最後，本書將展望人工智慧的未來發展，並探討如何在科技的發展和人類的福祉之間取得平衡。本書分為基礎理論和各類應用兩部分，基礎理論包含人工智慧、機器學習和深度學習，各類應用包括人文與法律應用、醫療應用、金融應用、行銷與零售應用、工業應用、農業應用以及永續、元宇宙和 AI 產業化進階的未來展望。

看完此書您一定會有 AI 的任督二脈被打通的舒暢感覺，看官不妨屏氣凝神好好邁向 AI 的新時代。

謝邦昌

天主教輔仁大學副校長

輔大 AI 人工智慧發展中心主任

台灣人工智慧發展學會 (TIAI) 理事長

中華市場研究協會 (CMRS) 榮譽理事長

中華資料採礦協會 (CDMS) 榮譽理事長

FOREWORD

推薦序

這是一本國內少有針對人工智慧教學的科普著作。適合給入門的讀者提供基礎的人工智慧知識，也適合作為學校科普教學用的書籍。作者累積多年在人工智慧的產業實務經驗，希望藉由人工智慧理論基礎及產業應用，讓讀者可以在短時間內熟悉此一領域。讓人工智慧知識不再是高不可攀，可以飛入平常百姓之家。對於剛入門的讀者可以先從第四章的實務應用開始閱讀到第九章，再回頭看前面有關人工智慧的理論基礎，學習會更有趣，收穫也會多更快。對於此書，因為有系統地介紹人工智慧，本人也誠摯地推薦可做為科普教學之用。

張順教

臺灣科技大學企管系教授

FOREWORD

推薦序

蔡宗翰教授是一位充滿熱誠，興趣廣泛的電腦科學家，曾聽他在演講中分享，從小就喜愛閱讀，內容遍及科學、歷史、人文、財經等各領域。他國中一年級被學校派去參加程式營，從此與資訊結下不解之緣，高中就讀建中科學班，但總是會找到一些喜歡文史同學來發掘相關的電腦應用。他在台灣大學資訊工程學系攻讀學士至博士學位期間，一路專注研究人工智慧，現在他分別在中央研究院和中央大學擔任研究員及教授，獲獎無數。只要翻開他的 Facebook，我們就能看到他深入淺出的歷史事件分析，見解都十分獨特新穎。目前他在「歷史+AI」及「中文+AI」兩方面的研究非常傑出，充分結合了他從小的興趣和長大之後的專長。

近年來，蔡教授專注於人工智慧教育，他所著的《寫給中學生看的 AI 課》更是名列博客來網站暢銷書第二名。由於使命感的驅動，他在繁忙的研究工作中，仍然奮力地和裴有恆老師完成這本《AI+AIoT 概論：寫給大學生看的 AI 通識學習》，將最新的 AI 科技和案例以淺顯易懂的方式傳播給讀者。由於我曾撰寫過 MATLAB 的教學書籍，我深知這是一項艱辛的工作。他們的貢獻使得這本書成為讀者理解和適應這種變革的重要工具，也讓我們更加明白 AI 和 AIoT 的潛力和可能性。

這本書的第一至第三章介紹了人工智慧、機器學習及深度學習，可以讓門外漢在不需要任何數學基礎的情況下，得以一窺 AI 的歷史、技術原理、應用、侷限等重要議題，同時也瞭解機器學習及深度學習的各項重要模型。而第四至第十章，則是重點式地介紹在各個領域的典型應用，包含人文與法律、醫療、金融、行銷零售、工業、農業等，可以讓各領域的專業人士瞭解 AI 及 AIOT 在特定領域的應用方式與目前的進程。

蔡宗翰教授和裴有恆老師的合作，將引領我們進入 AI 和 AIoT 的世界，幫助我們理解它們如何塑造我們的未來，讓我們一同期待並擁抱 AI 所帶來的變革，並以此為契機，開創更加繽紛多彩的未來。

張智星

玉山金控科技長

國立臺灣大學資訊工程學系 / 資訊網路與多媒體研究所教授

推薦序

感謝蔡宗翰教授與裴有恆理事長編撰此書，可用於通識教育學習人工智慧的基礎教材，提供了全面性的學理介紹與產業應用。

人工智慧（Artificial Intelligence，AI）是當今科技界最引人注目的熱門話題之一，這種技術不僅已經在多個產業領域得到廣泛的應用，而且正在推動整個社會的變革，進入到我們的日常生活之中。本書介紹了人工智慧的基礎理論到應用實例，將複雜的概念講解得淺顯易懂，並全方位地提供了生動的案例和實踐，為讀者展現了人工智慧的無限潛力。

本書的章節安排分為原理基礎和各種應用兩個部分。在原理基礎部分，作者從人工智慧、機器學習、深度學習等方面講解了各種演算法的運作原理，讓讀者對人工智慧的技術有深刻的理解。在各種應用部分，作者從人文與法律、醫療、金融、行銷零售、工業、農業、永續等多個領域中，挑選了代表性的案例進行了分析和解讀，讓讀者了解人工智慧在不同領域中的應用現狀和未來趨勢，同時也呈現出各行業導入人工智慧可能面臨的限制與挑戰。尤其是道德倫理、隱私、安全、歧視和就業市場變革等這些敏感議題，作者也提出了一些觀點，給讀者帶來了深刻的啟示。

在人工智慧浪潮下，近日有許多話題圍繞著，人工智慧可能會取代哪些職業或工作？科技是帶來人類社會的進步還是邁向滅亡？又或人工智慧也帶來了許多倫理道德與隱私安全的焦慮等等。其實，本書的作者也給了些許提示「某些人的工作內容可能會被生成式 AI 完全取代，而其他功能更有可能從人類和機器之間的密集迭代創意週期中獲益。科學家們對生成式 AI 的期許是將創造和知識工作的邊際成本降至零，從而產生巨大的勞動生產率和經濟價值。」事實上，近日 ChatGPT 的

推出襲捲全球，OpenAI 執行長 Sam Altman 在接受訪問時提到：「OpenAI 的目標是創造一種新秩序，在這秩序中，機器可以讓人們更充裕地做更多具創造性的事情。」國科會主委吳政忠也表示「AI 不會取代人類，但不會 AI 的人類會被取代」。由此可知，其真正的關鍵並不在於誰取代誰、或誰被誰取代，而是試著去認清 AI 是一種工具，能提昇個人與組織工作生產力的一種媒介。而科技所帶來的問題，或許也僅能靠科技的進步來逐步解決。

未來的組織團隊中，除了人之外，還有機器人，「人機協作」的教與學，是為必經之路。透過 AI 的使用，將人的思考腦袋、雙手雙腳、時間空間等從忙碌的工作與學習環境中空閒了出來，能夠做真正有價值的事情。換言之，AI 試圖引出人的潛力 (potential)，將人置放於更能發揮潛力的生活環境中。透過本書，對 AI 基礎原理與各行各業應用做個通盤性的瞭解，釐清各地的需求與挑戰，才能真正地「定位」您能發揮潛力與創造價值的地方。

在台灣正值推動「產業 AI 化」與「AI 產業化」的政策方向，協助 AI 人才培育與產業發展是主要的核心任務，相信本書對於初步踏入人工智慧領域的學習者來說，能帶來全面且通盤的視野。

李家岩

國立臺灣大學資訊管理學系教授

癸卯，仲夏

FOREWORD

推薦序

文組人在 AI 時代的必備指南

這本書是文組人必讀的一本書。

過去在中央研究院的網站、書店的暢銷書架、圖書館公開講座的海報，以及臉書的粉絲專頁上，常可看見 AI 李白—蔡宗翰老師的大名，知道蔡老師是一位資工出身又兼具人文素養的優秀學者。直到今年蔡老師當選臺灣數位人文學會理事長，在前輩的引薦下，才有機會正式深談。席間蔡老師以歷史為喻，談到 AI 將對人類產生的影響，心裡深深感佩蔡老師具有「樂以天下、憂以天下」的仁人志士情懷。後來得知蔡老師寫給大學生看的 AI 通識學習新書即將出版，希望人文背景的我幫忙寫推薦序。我二話不說的答應，除了滿足粉絲搶先看的私心，也樂於成為蔡老師推廣 AI 科普教育的一點助力。

我是一位中文系出身的人，接受的學術訓練包括文字、聲韻、訓詁、文獻、修辭。當年為了蒐集論文研究古籍資料，跑到各大圖書館一字一句輸入電腦。取得學位後，碰到古籍全文數據公開的熱潮，發現過去自己花一個月打字的古籍，文字辨識技術幾分鐘就可完成，實在有「字字看來皆是血，十年辛苦不尋常」的感嘆。痛定思痛，開始惡補各種數位人文研究方法。在跨領域閱讀過程中，很多資工專業詞彙像「有字天書」。遇見瓶頸時，就覺得如果有個理工人能幫文組的我們翻譯一下該有多好啊！幾年過去，陸續做出一些成果，看到人機合作的廣大世界；去年寫了《超數位讀國學》，希望邀集更多文組人勇敢跨出舒適圈。如今看到 AI 李白的登高一呼，希望能引起更多人重視數位知能。

什麼是監督式學習？什麼是多模態深度學習？卷積神經網路又是什麼？現在看不懂沒關係，AI 李白老師會用深入淺出的文字、舉例和圖像讓你理解。我想「絕對文組」的我都看得懂，正在翻閱此書的你，只要跟著李白老師文字和圖解的階梯，就能走進 AI 的世界。

過去文組人花費了很多時間在文獻整理，做的是「工人智慧」；在 AI 人工智慧的幫助之下，我們可以優化研究流程，不只事半功倍，甚至是事什功百。人文學在 AI 時代的價值因為不可取代的獨特性被標榜，理解 AI 是必經的過程。文組的我們不必學會寫程式，但必須理解演算法邏輯，發想問題，跨領域團隊合作，有效溝通，精煉人文價值。「以史為鏡，可以知興替」，活字印刷造成媒介革命，引發後續文藝復興、宗教改革、啟蒙時代和科學革命。在 AI 人機合作的數位時代，您我都不能置身事外。因時制宜，讓我們一起入門 AI，迎接新時代的來臨！

邱詩雯

國立臺灣師範大學華語文教學系助理教授
兼國語教學中心研發組組長

作者序

親愛的讀者，

在 2020 年的時候，因為台灣人工智慧學校昇瑋執行長的突然去世，我和裴老師來到了我在中央大學的辦公室相會，因為我們的文章都展現出對昇瑋的深深感念。當時我們並不熟。但那次見面，我們聊得很開心。我們才發現原來我們是學弟學長的關係，我一直在台大資工做 AI 研究，而裴老師在台大機械畢業後赴美深造，也在研究 AI。我們都認為，昇瑋的人工智慧普及教育理念必須繼續推進。

接下來的兩年，我出版了《寫給中學生看的 AI 課》這本書，也在博客來資訊類和親子教養類排行榜上拿到了不錯的名次。而裴老師也在 AIoT 領域發表了一系列受到熱烈歡迎的書籍。我們都在自己的領域中不懈努力，期待能對台灣的 AI 教育做出貢獻。

2022 年底，當我正忙著寫《寫給小學生看的 AI 課》這本書的時候，裴老師打來電話，邀請我和他共同創作一本針對大學生的 AI+AIoT 通識讀本。我負責 AI 基礎原理部分，他負責 AI+AIoT 的應用部分。我當時擔心自己沒時間，不過最後還是被裴老師的熱誠說服，決定一起寫這本書。

回頭看看我自己的大學生涯，我第一次接觸 AI 就是在大一的時候，那時候我在許永真教授的計算機概論課中寫了一支 AI 程式。那時的研究過程讓我充滿成就感，也為我未來在 AI 領域的研究埋下了種子。

但學習 AI 並不容易。您可能會遇到很多困難，可能會覺得自己不能理解，甚至可能會想放棄。但我希望您能夠堅持下去。因為 AI 的魅力不在於它的高深複雜，而在於它如何能解決我們生活中的問題，如何能提升我們的生活品質。

所以我們決定寫這本《AI+AIoT 概論：寫給大學生看的 AI 通識學習》，希望它能像一本指南，帶領您在 AI 的世界中尋找方向，也讓您了解到 AI 和 AIoT 的應用可以如何影響我們的生活。我們還希望，透過這本書，讓大學生有更多的機會接觸 AI，也讓台灣的 AI 教育能更上一層樓。

寫這本書的過程非常愉快，也非常有挑戰，但我們都非常滿意我們的成果。我們希望這本書能對您有所幫助，也希望您能享受學習 AI 的過程。歡迎大家追蹤我的 fb 粉絲團「AI 界李白」，持續收到最新的 AI 新知與資源。有任何問題或邀約也可以透過這個管道。

祝學習愉快！

AI 界李白 蔡宗翰 老師



AI 界李白 FB

作者序

2016年，我推出第一本物聯網商機的書籍：《改變世界的力量、台灣物聯網大商機》；2017年，在碁峯資訊的協助下，完成了第二本書：《物聯網無限商機—產業概論 x 實務應用》；2018年，完成了《AIoT 人工智慧在物聯網的應用與商機》第一版。這是我的第三本書，它還出到了第三版。2019年之後，我每年出一本 AIoT 數位轉型的書籍，在2022年，我在碁峯資訊出版了我的第七本書《從穿戴運動健康到元宇宙，個人化的 AIoT 數位轉型》。

就如 AI 界李白蔡宗翰老師在他的作者序所言，我們在 2020 年台灣人工智慧學校陳昇璋執行長突然過世的時候，因為感佩這樣一位 AI 界好戰友的過世而相聚，並訂下了發揚昇璋遺願，讓人工智慧教育普及的同志願。而在 2022 年，碁峰資訊跟我熟稔的產品企劃發現坊間沒有適合大學用的 AI 通識書籍，因為多年的合作，她便找上我希望我能協助，我就想起之前跟蔡老師談的共同推廣人工智慧教育普及的志願，認為如果有蔡老師在 AI 基礎理論的深耕，以及在人文及醫學方面 AI 多年的專案經驗，再加上蔡老師寫了《寫給中學生看的 AI 課》這本對中學生很棒的科普書籍，結合我在 AI 及 AIoT 各應用領域的多年深耕，兩個人合作出書一定可以讓大學生對 AI 原理與各類應用，有清楚而適合的了解，也因此我在跟蔡老師一起規劃課程內容時，除了前三章的基礎 AI 理論，在應用面考慮到各產業的需求，從第四章到第九章，針對人文與法律、醫療、金融、零售、工業、農業來談各種 AI 或 AIoT 的應用案例，而第十章，就針對 AI 的未來展望來談對 ESG、元宇宙，以及產業 AI 化進階來討論發展可能性。畢竟，AI 的浪潮來勢洶洶，特別是在 ChatGPT 廣被應用之後，相信大家更能深刻感受到，現代人如果不學 AI，很快就會被時代淘汰掉，而我們合著的這本科普書籍《AI+AIoT 概論：寫給大學生看的 AI 通識學習》，正是幫助帶學生及各行各業的人對 AI 原理及各類應用有清晰而完整的科普知識。

這本書能夠出版，首先要感謝蔡宗翰老師願意在百忙中撥冗完成相關的內容部分，接著特別要感謝我的博士班指導教授張順教教授、輔仁大學謝邦昌副校長、台大資管系李家岩教授，以及台大資工系教授暨玉山金控科技長張智星老師願意協助提供推薦序，也非常感謝全家便利商店數位轉型部副本部長林翠娟、台灣金融科技協會監事林玲如、台灣大學農藝學系劉力瑜教授、臺北護理健康大學語言治療與聽力學系翁仕明副教授、國立臺灣師範大學華語文教學系邱詩雯助理教授，以及台大創新領域學士學位學程袁千雯專任副教授的願意推薦此書，這也表明了這本書是為了各行業人了解 AI 的科普書籍。

也歡迎各位讀者透過 Google 查詢「Rich 老師的創新天堂」找到部落格、臉書粉絲專頁或我的公司「昱創企管」官網跟我聯繫。如果想開始了解如何用 AIoT 完成數位轉型，也歡迎購買我出的《白話 AIoT 數位轉型》、《AIoT 數位轉型策略與實務》，兩本書以做更深入的了解。而針對製造業，另外，我也跟獲得磐石獎的新呈工業董事長陳泳睿合出了《AIoT 數位轉型在中小製造企業的實踐》這本講述智慧工業各國標準，加上台灣中小製造業案例的書籍，這是讓製造業的中小企業對 AIoT 數位轉型有深入的概念。另外我的 YouTube 頻道「數智創新力」可以幫助大家對 AIoT 綠色轉型及數位轉型建立基礎概念，歡迎大家去訪問。

裴有恆 Rich

中華亞太智慧物聯發展協會創會理事長

好食好事基金會業師

臉書社團：i 聯網、智慧健康與醫療創辦者

昱創企管顧問有限公司總經理



i 聯網



昱創企管顧問
有限公司



Rich 老師的
創新天堂

INTRODUCTION TO
AI+AIOT



PART

I

AI 的原理基礎

第 1 章 人工智慧

第 2 章 機器學習

第 3 章 深度學習

本書的前三章將介紹人工智慧（AI）的原理基礎。第一章將從概念和應用方面介紹 AI，並探討其局限性和道德倫理問題。第二章將深入介紹機器學習（ML）的基本概念和技術，以及在現實世界中的應用和發展趨勢。第三章將探討深度學習（DL）的基本知識和在解決複雜問題上的應用。透過這三章的學習，讀者可以獲得一個全面的 AI 技術概觀，以便更清楚其在不同領域的應用和未來發展趨勢。

✔ 第一章：人工智慧

本章介紹人工智慧的概念、類型，並透過實例說明技術和應用，探討其局限和挑戰。同時也討論人工智慧的道德和倫理問題，包括透明度、偏見和就業影響等，最後展望未來發展和平衡科技和人類福祉。

✔ 第二章：機器學習

本章深入介紹機器學習的基礎概念，包括不同演算法、關鍵詞彙和技術。同時也舉例展示機器學習在現實中的應用和對社會造成的影響。

✔ 第三章：深度學習

深度學習是機器學習領域裡的一個分支，透過訓練神經網路自主學習和做出決策，在圖像、語音、自然語言等方面取得了不少傲人的成果。這一章會深入探討深度學習的基本概念，以及在解決各種複雜問題上的實際應用。

大家可以針對自己有興趣的部分開始閱讀吧！

1 人工智慧

在這一章中，我們將認識人工智慧的概念、類型，透過舉例來解釋人工智慧的各種技術和應用，以及探討人工智慧的局限性和潛在挑戰。接著，我們也將討論人工智慧的道德與倫理問題，包括人工智慧決策的透明度和可追溯性、偏見和歧視、對就業市場的影響等。最後，我們將展望人工智慧的未來發展方向，並探討如何在科技發展和人類福祉之間取得平衡。

1.1 什麼是人工智慧 (Artificial Intelligence) ?

創造能夠「思考與行動」的智慧型機器，開發能分析和理解資料、從經驗中學習並根據知識做出決策的演算法和系統，是人工智慧科學家的共同目標。簡單來說，人工智慧是一門「研究如何使機器具有『像人類一樣』的智慧」的學科。

自萌芽以來，人工智慧科學已經歷了顯著的發展和演變，因此定義也十分多元，人工智慧領域先驅者比較廣為人知的定義有以下幾種：

| 定義者 | 定義內容 |
|------------------------|-----------------------|
| 艾倫·圖靈 (Alan Turing) | 機器展現類似於人類的智慧行為的能力。 |
| 約翰·麥卡錫 (John McCarthy) | |
| 大衛·哈雷爾 (David Harel) | 研究如何使電腦做「人現在勝過電腦」的事情。 |



| 定義者 | 定義內容 |
|--------------------------|-----------------------------------|
| 約翰·麥卡錫 (John McCarthy) | 使電腦表現得像人類的電腦科學分支。 |
| 斯圖爾特·羅素 (Stuart Russell) | 創建智慧型機器（尤其是智慧電腦程式）的科學和工程。 |
| 彼得·諾維格 (Peter Norvig) | |
| 馬文·明斯基 (Marvin Minsky) | 使機器做人類需要智慧才能完成的事情的科學。 |
| 吳恩達 (Andrew Ng) | 致力於創建能夠推理、學習和自主行動的智慧代理的電腦科學和工程領域。 |

除了艾倫·圖靈 (Alan Turing) 和約翰·麥卡錫 (John McCarthy) 最早提出的定義，後來有些學者的定義較為廣泛，涵蓋了**機器完成需要人類智慧才能完成的任務的能力**，本書則把人工智慧定義為**研發能夠執行需要人類等級智慧的任務（例如學習、解決問題和做決策）的電腦系統**，包含各種人工智慧技術，如機器學習、自然語言處理和電腦視覺……等。其目標在於使機器擁有「做出人類智慧所能及之事」的能力，例如做決策、解決問題、學習、解謎、對話、辨識物體、辨識聲音、辨識圖像等。一言以蔽之，人工智慧就是**幫助人做事**的機器。

如今，人工智慧雖然還沒辦法剖析所有人類智慧，但人工智慧已肩負起「幫助人做事」的任務，並且正在改變我們的生活和工作方式。例如將任務自動化以使任務執行更有效率，以及提供新的見解和解決複雜問題的方案。隨著這一領域的不斷發展，仔細探究人工智慧對道德和社會影響，並確保以負責任和透明的方式研發和使用人工智慧，將至關重要。

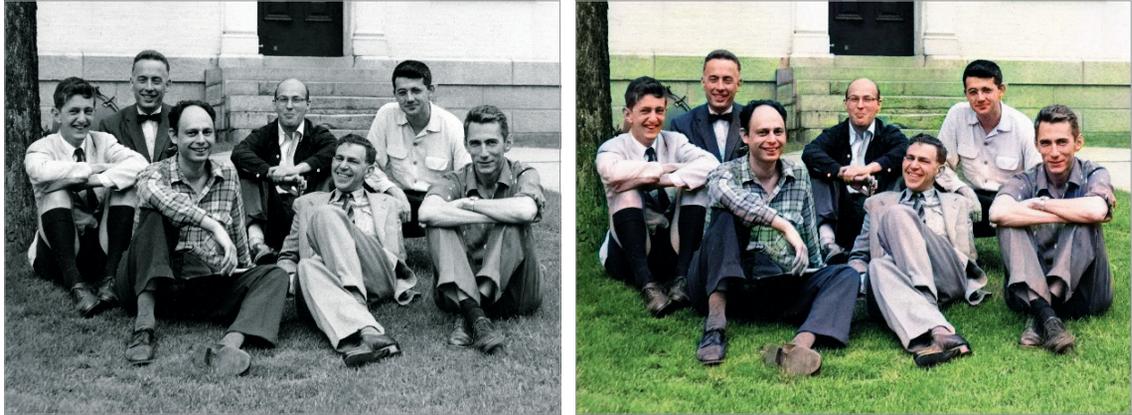
1.2 人工智慧的歷史

初聲

約在 1940 年代，研究者們萌生了創建智慧型機器的想法。著名的資訊科學家艾倫·圖靈在 1950 年代提出了一個問題：

「能否設計一種電腦程式，讓它能像人類一樣思考？」這就是人工智慧的起源。

1956年夏天，人工智慧領域最早也最有影響力的事件之一——「**達特茅斯會議** (Dartmouth Summer Research Project on Artificial Intelligence)」在美國的達特茅斯學院舉行。這次會議聚集了包括約翰·麥卡錫、馬文·明斯基在內的一群頂尖研究人員（下圖左為原合照，右為 AI 將左圖自動上色的結果），期間，這群科學家熱烈地討論了人工智慧的發展潛力，而「人工智慧 (Artificial Intelligence)」一詞更在這次會議正式誕生。



圖片來源：Photograph taken by Gloria Minsky

1950 和 60 年代

人工智慧研究集中在創建能夠執行特定任務的程式，例如下棋或解決數學問題。這些早期人工智慧程式，稱為「專家系統」，在能力方面有限，只能在限定的專業領域內執行任務。在 1970 和 80 年代，研究人員開始致力於開發能夠學習和適應新情況的人工智慧系統。這導致機器學習演算法的發展，這些演算法可以分析資料並提高其性能。在這段時間，人工智慧研究也開始分支出許多不同的次領域，包括自然語言處理、電腦視覺和機器人技術。這些次領域繼續發展並變得越來越重要，並對許多產業產生了重大影響。

1990 年代和 2000 年代

人工智慧研究再度興起，部分原因是電腦硬體的進步和大量資料的可用性。這一時期出現了深度學習演算法，這些演算法已被用於在圖像和語音辨識等許多領域實現最先進的性能。此期間的一個關鍵事件是 ImageNet 的出現，這是一個大型圖像資料庫，用於訓練和評估機器學習演算法。



2010 年代初期

人工智慧領域取得了重大進展，開發了新的演算法和技術，對各種應用產生了重大影響。

其中一個關鍵發展是使用「預訓練模型」。這些模型是已經使用大量資料訓練過的機器學習模型，可以微調使用在特定任務上。

2017 年以來

近年來，人工智慧領域最重要的發展之一是 2017 年「Transformer」的出現。這種神經網路架構，徹底改變了自然語言處理（NLP）領域，更被應用於電腦視覺和音樂處理，顯著推進了這些領域的研究。另一個獲得注意的新興研究領域則是使用大量未標註資料訓練機器學習模型的「自監督學習」，這種方法已被用於在圖像和語音辨識¹等許多領域上，並展現了最先進的性能。

總之，過去十年機器學習演算法和技術的發展，使人工智慧在各領域應用的廣度與深度都有所突破，不過人工智慧的潛力遠大於現今我們所看到的。人類仍有很多研究工作要做，才能完全理解和利用它的潛力。

1.3 人工智慧的類型

人工智慧可以依不同的角度分類，主流區分方式有二：依能力區分、依目的區分。

若按照**能力**分類，可分為三種類型：弱人工智慧、強人工智慧和專家系統。

✔ 弱人工智慧 (Weak AI)

也稱為「特定型人工智慧」。弱人工智慧並不具有人類般的智慧，而是專注於解決某特定問題的能力。例如，智慧客服系統就是一個弱人工智慧的例子，它專門用於回答顧客的問題，像是家電故障排解、預約特定服務等，但無法與顧客聊天或整合資訊。現今所見到人工智慧系統幾乎都屬於弱人工智慧。

1 Yang, S. W., Chi, P. H., Chuang, Y. S., Lai, C. I. J., Lakhotia, K., Lin, Y. Y., ... & Lee, H. Y. (2021). Superb: Speech processing universal performance benchmark. Interspeech 2021.



1.4 人工智慧的技術

常見的人工智慧的技術有下列四種：

✓ 機器學習 (Machine Learning)

讓機器透過學習經驗來改善自己的表現。機器學習分為監督學習、非監督學習、半監督學習和增強學習四種：

- ⊗ 監督學習是指機器透過訓練資料學習如何解決問題，訓練資料包含了「輸入資料」和「對應的正確答案」。
- ⊗ 非監督學習是指機器透過分析大量的資料來發現資料之間的關聯，而無須給定正確答案。
- ⊗ 半監督學習是指機器同時使用有標註資料和無標註資料來學習。
- ⊗ 增強學習是指機器透過反覆嘗試和獲得經驗來改善自己的表現。

✓ 自然語言處理 (Natural Language Processing, NLP)

用於電腦能夠理解人類的自然語言，並能夠使用自然語言²和人類互動。這個領域中還有「自然語言理解」和「自然語言生成」兩個子任務：

- ⊗ 自然語言理解 (Natural Language Understanding, NLU) 透過分析語法、語義、情感和其他資訊來讓電腦能夠理解人類自然語言的意義。
- ⊗ 自然語言生成 (Natural Language Generation, NLG) 透過分析數據、撰寫報告或生成回應等方式，讓電腦能夠使用人類自然語言生成文本。

NLP 技術可以應用許多非常仰賴語言的領域，如聊天機器人、語音辨識、翻譯、文本摘要、情感分析、自然語言查詢等。

2 自然語言是指人類之間溝通所使用的語言。

✔ 影像處理 (Image Processing)

用於處理和分析圖像和影像數據。在影像處理中，常見的任務包括圖像分類、圖像分割、物體識別、圖像生成和圖像風格轉換：

- ③ 「圖像分類」是將圖像分類為不同的類別的過程，通常是通過訓練一個模型將圖像的特徵與類別標註相關聯。如 AI CUP 2022 「蘭花種類辨識與分類競賽」³，就是一個典型的圖像分類任務。
- ③ 「圖像分割」則是將圖像中的不同區域分類為不同的類別，通常是通過訓練一個模型將圖像中的每個圖元分類為特定的類別。如 AI CUP 2022 「肺腺癌病理切片影像之腫瘤氣道擴散偵測競賽 II：運用影像分割做法於切割 STAS 輪廓」⁴，就是一個典型的圖像分割任務。
- ③ 「物體識別」是在圖像中檢測和定位特定物體的过程，通常是通過訓練一個模型可以在圖像中檢測出特定物體，並給出每個物體的位置。如 AI CUP 2022 「肺腺癌病理切片影像之腫瘤氣道擴散偵測競賽 I：運用物體偵測做法於找尋 STAS」⁵，就是一個典型的圖像分割任務。
- ③ 「圖像生成」則是使用人工智慧技術生成新圖像的過程，通常是訓練一個模型根據給定的模式或條件生成新圖像。
- ③ 「圖像風格轉換」是將圖像的風格轉換為另一種風格的過程，通常是通過訓練一個模型可以將輸入圖像的風格轉換為另一種指定的風格。

此外，還有許多相關的輔助型任務，例如圖像品質評估、圖像去噪、圖像增強等。

✔ 語音處理 (Speech Processing)

用於處理、分析、合成和辨識語音訊號。其中，「語音辨識 (Speech Recognition)」可以讓機器辨識人類語音並將其轉換成文字，「語音合成 (Speech Synthesis)」則可以讓機器將文字轉換成語音訊號，並通過喇叭或耳機等輸出。這兩項技術已經在手機、語音 / 翻譯軟體、汽車和家庭自動化系統隨處可見，可說是現階段最貼近人類生活的人工智慧技術之一。

3 <https://www.aicup.tw/ai-cup-2022>

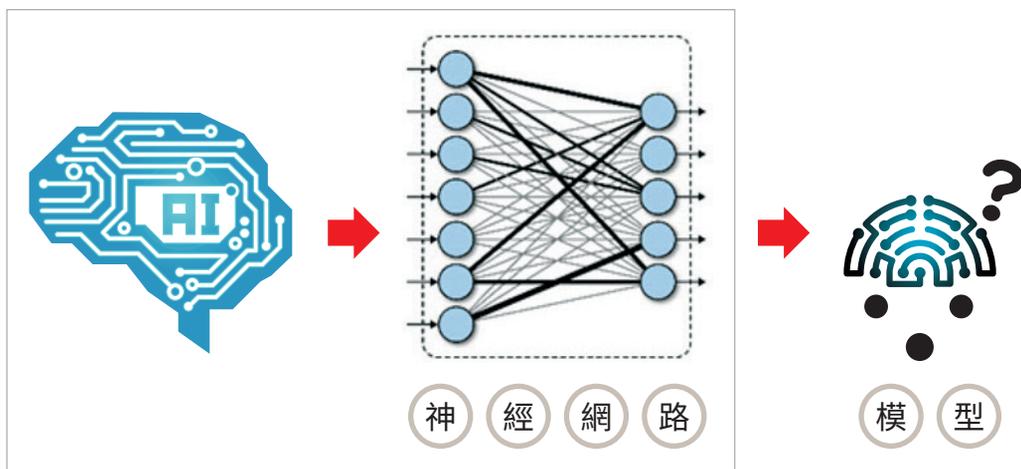
4 <https://www.aicup.tw/ai-cup-2022>

5 <https://www.aicup.tw/ai-cup-2022>

3 深度學習

深度學習是機器學習的一個次領域。機器受到人類大腦結構和功能的啟發，尤其是人類大腦中神經網路的運作方式，促使它能夠透過大量資料集來訓練神經網路，讓神經網路能夠自主學習並做出有智慧的決策。在各種不同領域中，深度學習都取得了領先的成果，包括圖像和語音辨識、自然語言處理，甚至是自動玩遊戲。本章我們將深入探討深度學習的基本知識，以及它如何被運用在解決複雜的問題。

為了幫助沒有深度學習操作經驗的讀者易於理解，後文描述時，我們多採取下圖的架構來說明：執行任務的智慧體統稱為 AI，AI 所依據的神經網路架構稱為神經網路或網路，由訓練得到神經網路的參數資料稱為模型。AI 在進行推理時需要透過神經網路，而神經網路運作時需要載入模型中的參數資料。





3.1 深度學習的基礎知識 (Loss function)

深度學習 AI 可以看成是一個黑箱子。與機器學習一樣，深度學習 AI 的運作可以分成訓練階段及測試階段（或稱推論階段、使用階段），不論是訓練階段或是測試階段，它都必須接受一筆資料實例做為輸入，然後會輸出該筆資料實例對應的標籤。例如，我們可以輸入一張圖，經過 AI 的網路，它就會輸出圖的標籤（標籤的類別必須事先定義好）。在訓練階段，AI 尚未成熟，因此輸出的標籤機率分布會與真正的標籤機率分布做比對，以損失函數 (Loss function) 計算出這次預測的損失，用來反饋、修正神經網路內部的權重。在測試階段，我們則是會直接採用 AI 預測的結果，並不會去計算損失。



損失函數

損失函數 (Loss function) 是在機器學習中用來評估「預測結果」與「實際結果」之間的差距的函數。而訓練 AI 過程中的目標是最小化損失函數的值。

各種任務有其適合的損失函數，因此，AI 設計者必須做出選擇。例如，假設我們要訓練一個線性迴歸 AI 以預測房屋價格。在訓練過程中，我們可以使用均方誤差 (Mean Squared Error, MSE) 作為損失函數，它可以幫助我們評估預測值與實際值之間的差距。此外，還有許多其他的損失函數，例如，交叉熵 (Cross-entropy) 損失函數常用於分類問題。接著我們就用交叉熵來舉例說明 AI 如何計算一次分類預測的損失，計算交叉熵損失函數必須用到「預測的各類別的機率」和「實際類別的機率」。

例如，假設我們有一個圖像分類模型，預測一張圖片是狗的機率是 0.8，是貓的機率是 0.2，但實際上這張圖片是狗。我們可以計算交叉熵損失函數：

$$-loss = -(y \times \log(p)) + (1 - y) \times \log(1 - p)$$

$y = 1$ 代表實際類別是狗， $p = 0.8$ 代表模型預測這張圖片是狗的機率為 0.8。

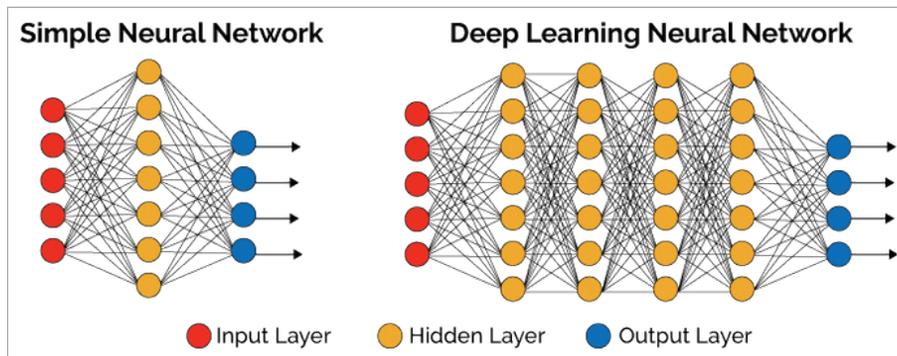
$$\begin{aligned} -loss &= -(1 \times \log(0.8)) + (1 - 1) \times \log(1 - 0.8) \\ &= -(0.22314) = 0.22314 \end{aligned}$$

這就是這次圖片預測的交叉熵損失函數值。

深度學習的資料實例表現形式

為了由淺入深，我們前面並沒有提到精確的輸入資料形式。實際上，在深度學習中，資料實例通常以數值陣列的形式表現，稱為「向量」。這些向量會被輸入神經網路，該網路由互相連接的節點層組成，稱為「神經元」。

每一層的神經網路都會處理輸入的資料，並將這些資料傳遞到下一層，直到最終層產生輸出預測（如下圖）。輸入和輸出層之間的層稱為隱藏層，因為它們的內部工作是用戶不可見的。下圖左方的是最簡單的神經網路，共有三層；超過三層以上的就稱為深度神經網路（如下圖右）。目前，深度學習所指的就是 AI 透過訓練過程，學會深度神經網路中的所有權重。



圖片來源：<https://towardsdatascience.com/mnist-vs-mnist-how-i-was-able-to-speed-up-my-deep-learning-11c0787e6935>

深度學習的訓練過程

如同前面所述，AI 設計者必須針對不同任務，挑選損失函數與優化算法。損失函數用來衡量 AI 在訓練資料上的表現，而優化算法則可調整 AI 使用的神經網路中的權重與偏差，以最小化損失。

開始訓練 AI 後，我們將訓練資料批次輸入，並使用損失函數計算損失。然後，再使用優化算法調整權重和偏差，以減少損失。此過程重複進行多次，AI 不斷學習並改善其對訓練資料的預測。

等到訓練完成，我們就可以在另一個測試資料集上評估 AI 的性能，以查看它對新資料的泛化能力如何。如果表現良好，我們就可以使用它來預測新的資料實例。

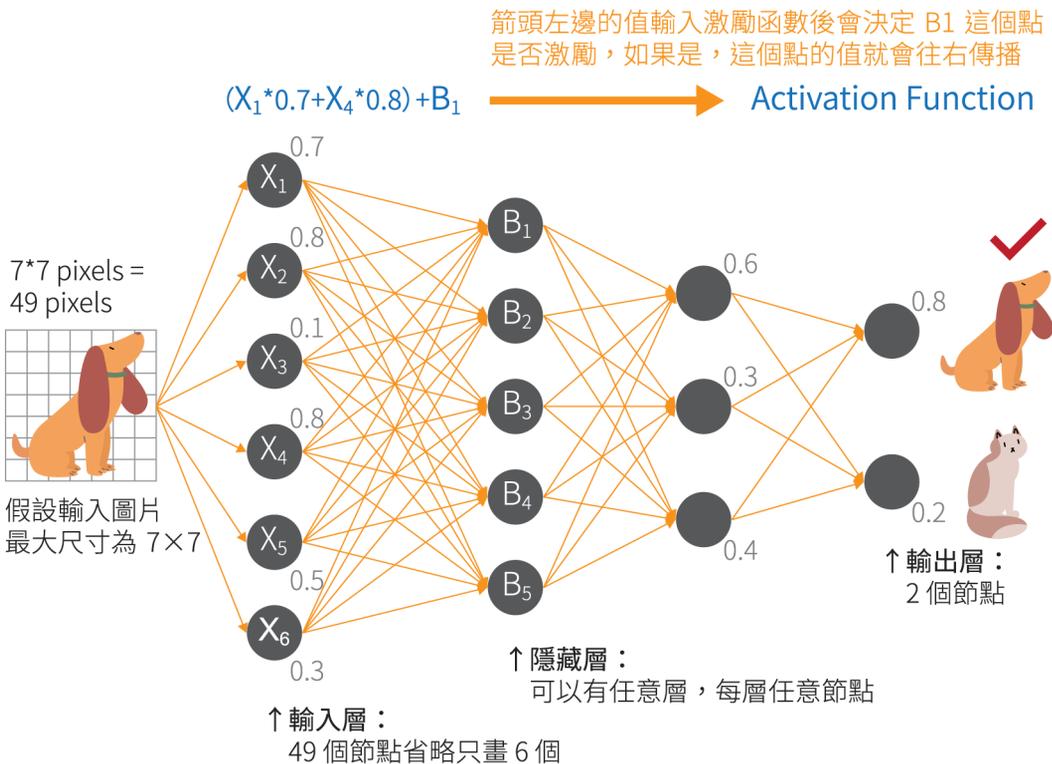


深度學習的主要優點是能對資料中的複雜模式進行學習與建模，這要歸功於神經網路的多層次結構。每一層都能夠擷取資料裡越來越抽象的特徵，使 AI 能夠做出更精確的預測。

總之，訓練深度學習 AI 的過程包括將資料以向量表現，針對各種任務選擇適當的損失函數和優化算法，並在訓練資料上迭代調整神經網路的權重和偏差，以最小化損失，這使得 AI 能夠自主學習並做出最理想的決策。

案例：圖像分類

假設我們想訓練一個深度學習 AI 來將狗和貓的圖像分類。首先，我們必須輸入狗和貓的圖像資料實例，用來表示向量。模型則是設定有兩個輸出類：「狗」和「貓」。接著，我們開始使用狗和貓的圖像資料集訓練 AI，其中每個圖像都被標記為「狗」或「貓」。假設 AI 能接受的圖片最大尺寸為 7×7(如下圖)，攤平後成為一個 49×1 的向量，也就是輸入層將有 49 個節點；中間可以有任意數量的隱藏層，每一層都可以有任意個節點；輸出層則有兩個節點，分別對應到狗與貓。訓練的過程中，AI 將學習把圖像中的某些特徵與相應的標籤作出聯結。例如，模型可能會學到長耳朵和蓬鬆的尾巴是貓的特徵，而短鼻子和尖耳朵是狗的特徵。



→ Forward Propagation

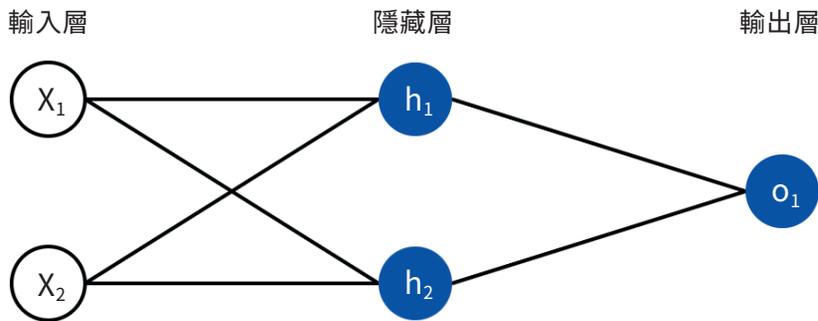
3.2 深度學習的基本神經網路類型

前饋神經網路 (Feedforward Neural Network, FNN)

前一節已經用某種神經網路來解釋深度學習的過程。事實上，這種神經網路稱為前饋神經網路 (Feedforward Neural Network, FNN)。它包括了輸入層 (input layer)，一個或多個隱藏層 (hidden layer) 和輸出層 (output layer)。輸入層接收輸入資料，隱藏層處理資料，輸出層產生最終輸出。

在傳統的 FNN 中，資訊只能從輸入層流向隱藏層，最終導向輸出層。因為沒有循環或回溯的連接，所以資訊無法流向相反的方向。

隱藏層及輸出層中的資訊處理是透過在訓練過程中學到的權重 (weight) 和偏差 (bias) 來完成的。權重用於縮放輸入資訊，偏差用於移動資訊，兩者合力將輸入資訊轉換為適合當下執行的任務。轉換之後，還需套用激勵函數 (activation function)¹，將轉換結果映射至 0 到 1 之間的區間。



讓我們以一個簡單的例子來說明 FNN 如何計算結果。我們有一個健康資料集，每筆資料有三個值：身高、體重、是否過重。我們想要利用這個健康資料集訓練一個基於 FNN 的 AI：輸入身高體重值，就能判斷此人是否過重。因此，輸入為二維向量， x_1 對應到身高， x_2 對應到體重。我們可以針對原始資料中的是否過重欄位值進行修改，「是」改為 1，「否」改為 0。假設對於每個節點，訓練中學習到的輸入層連接到隱藏層之權重 ((x_1, h_1) 表示 x_1 連到 h_1 的權重) (如下表)：

¹ 激勵函數 (activation function) 將輸入值映射至 0 到 1 之間的區間。常見的 activation function 有 tanh、ReLU、sigmoid 等等。



| | h_1 (偏差 0.2) | h_2 (偏差 0.1) |
|-------|-----------------|-----------------|
| x_1 | 0.4 | 0.7 |
| x_2 | 0.6 | 0.3 |

我們有：

$$h_1 = \text{activation}(0.4 \times x_1 + 0.6 \times x_2 + 0.2)$$

$$h_2 = \text{activation}(0.7 \times x_1 + 0.3 \times x_2 + 0.1)$$

而隱藏層連結到輸出層之權重（如下表）：

| | o_1 (偏差 0.15) |
|-------|------------------|
| h_1 | 0.8 |
| h_2 | 0.2 |

我們有：

$$o_1 = \text{activation}(0.8 \times h_1 + 0.2 \times h_2 + 0.15)$$

不論輸入值是多少，激勵函數都可以將輸入映射至 0 與 1 之間的區間。因此，若最終 $0.8 \times h_1 + 0.2 \times h_2 + 0.15$ 代入激勵函數的輸出值大於等於某個閾值，我們就可以說神經網路分類的結果為：此人過重。我們也可以更改 FNN，將輸出層改為兩個節點，節點 o_1 代表過重， o_2 代表未過重。如果 o_1 的值比 o_2 大，代表分類的結果為此人過重，反之則為此人未過重。

卷積神經網路 (Convolutional Neural Network, CNN)

由 FNN 處理圖像的例子，您可能會發現它的輸入層與圖像的像素數相等。如果一張 1000×500 的圖像，輸入層就會有高達 500,000 個節點。假設第一個隱藏層有 1000 個節點，光是輸入層到第一個隱藏層的權重參數就高達 500000×1000 這麼多！AI 要學習的參數數量過於龐大，也代表訓練用的資料與訓練用的運算資源都需要非常龐大，這使得 FNN 幾乎無法直接應用。而且，一張圖像裡面，有許多相似的地方。例如，這裡有一張狗的圖像（如下圖左），將圖像放到最大（如下圖右），可以

發現狗臉上的毛細孔由左到右重複出現。因此，像 FNN 這樣把每個像素都當作是一個不同的單元，對應到一個權重讓 AI 去學習，實在是太奢侈了！



卷積神經網路（Convolutional Neural Network，CNN）就是用來解決 FNN 在圖像上應用不易的問題。由於它是一種專門用於處理具有網格狀資料的神經網路，因此特別適用於圖像相關的任務，因為它可以學習辨識圖像中的模式和特徵。

CNN 的推論過程包括將輸入圖像經過一系列層，每層都使用一組學習過的權重矩陣，稱為濾波器（filter），對輸入的圖像進行運算。filter 由一到多個 kernel 組成。舉例來說，如果圖像是灰階，因為只有一個頻道，所以一個 filter 就只有一個 kernel；如果圖像是彩色，有 RGB 三個頻道，一個 filter 就會有三個 kernel。而每個 kernel 都是在訓練過程中學習出來的，在訓練的過程裡，將一組訓練圖像輸入 CNN 中，AI 對 CNN 中的權重進行調整，使得 CNN 能夠對圖像進行正確的分類。同時，filter 下所屬的每個 kernel 也會不斷地被調整，以檢測出圖像中最重要的特徵。

接下來，我們用一個例子來介紹在「圖像分類任務」中，AI 怎樣使用 CNN 進行推論。假定輸入圖像尺寸統一為 6×6 ，格式為灰階單頻道，此次輸入的圖像（如下圖左），0、1 分別表示該像素為黑、白；我們要辨識該圖像是否含有「尖角向左三角形」這個特徵。假定透過訓練資料，AI 已經學出了一個 3×3 的 kernel（如下圖右）。此時，因為一個 filter 只含有一個 kernel，所以 filter 就等於 kernel。

在計算卷積時，由輸入圖像的左上角 (1,1) 開始，將 kernel 疊上去，並計算 kernel 和輸入圖像之間的元素乘積。元素乘積的算法就是同位置的元素相乘，再整個加總起來。於是我們有：



$$0(-1) + 0(-1) + 1(1) + 0(-1) + 1(1) + 1(1) + 0(-1) + 0(-1) + 1(1) = 4$$

| | | | | | |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 |
| 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 |

| | | |
|----|----|---|
| -1 | -1 | 1 |
| -1 | 1 | 1 |
| -1 | -1 | 1 |

接著，根據設定的平移數 (stride)，假定為 1，往右移動到 (1,2)，將 kernel 疊上去，並計算 kernel 和輸入圖像之間的元素乘積。於是我們有：

$$0(-1) + 1(-1) + 0(1) + 1(-1) + 1(1) + 0(1) + 0(-1) + 1(-1) + 1(1) = -1$$

| | | | | | |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 |
| 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 |

| | | |
|----|----|---|
| -1 | -1 | 1 |
| -1 | 1 | 1 |
| -1 | -1 | 1 |

這樣依序往右移動到 (1,4)，就完成了第一列的計算，分別得到 4、-1、-3、1 四個值。我們可以用同樣的方式計算第 2-4 列，按照列數由上往下擺放，最後會得到一個矩陣 (如下圖)。

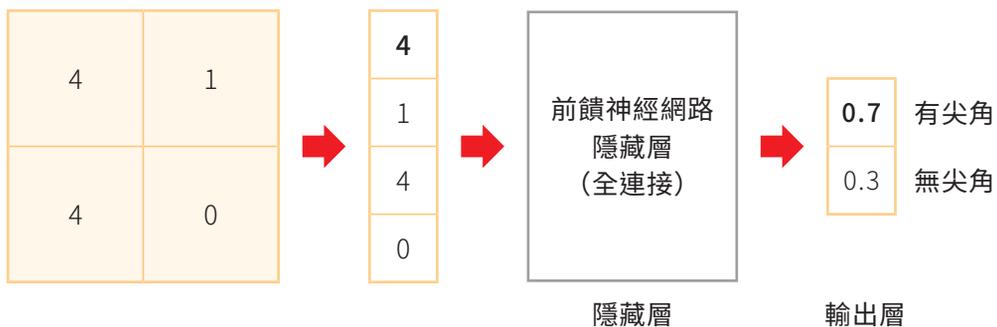
| | | | |
|----------|----|----|----|
| 4 | -1 | -3 | 1 |
| 2 | -1 | 0 | -3 |
| 2 | -1 | -2 | -1 |
| 4 | -2 | 0 | -1 |

我們稱這個矩陣為 feature map，表示 kernel 在輸入圖像的每個位置的反應。您是不是有發現， 6×6 的圖中，(1,1) 與 (4,1) 為左上角的 3×3 區域，出現了 kernel 所設定的尖角向左三角形特徵呢？

因為這個 feature map 還是太大，所以我們要採用池化 (pooling) 來減低它的大小。最常採用的池化方式就是最大池化 (max pooling)。假設我們採用 2×2 為 pooling 的尺寸，於是上面的 feature map 就會變成四個子區域 (以粉紅色標示，如下圖左)。四個子區域分別選出自己的最大值做為代表，得到的結果如下圖右：

| | | | | | |
|---|----|----|----|---|---|
| 4 | -1 | -3 | 1 | 4 | 1 |
| 2 | -1 | 0 | -3 | 4 | 0 |
| 2 | -1 | -2 | -1 | | |
| 4 | -2 | 0 | -1 | | |

到目前為止，透過卷積與池化，我們已經將輸入圖像的尺寸由 6×6 縮小到 2×2 (如下圖)。因為 2×2 已經夠小，我們就可以把 2×2 的矩陣攤平成為 4×1 的向量，接著連結前面介紹的 FNN，如同前面所述，這個 FNN 由一連串全連接層所構成，就可以推算出圖像屬於各類別的機率分布 (例如，「有尖角」與「無尖角」)，機率最高的類別將是 CNN 做出的最終預測。在下圖中，我們假定 CNN 預測圖像為有尖角與無尖角的機率依序是 0.7 與 0.3，因此這次預測的結果就是「有尖角」。

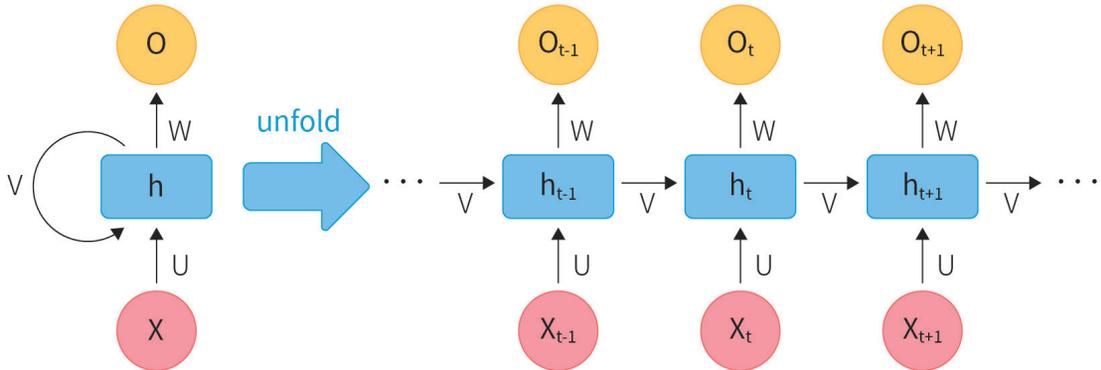


接著我們來比較用 CNN 可以比 FNN 節省多少參數。若 FNN 與 CNN 在特徵擷取後，同樣為含有 3 個節點的隱藏層，使用 FNN，輸入層連結到隱藏層的權重數量是 $(6 \times 6) \times 3 = 108$ ；但使用 CNN，輸入層連結到隱藏層的權重數量是 $(2 \times 2) \times 3 = 12$ ，僅為 FNN 的 $1/9$ 。由此可以印證 CNN 確實可以大幅減少神經網路的參數量。



循環神經網路 (Recurrent Neural Network, RNN)

循環神經網路 (RNN) 是一種特殊的神經網路架構，它能夠處理時間序列的資料。RNN 透過循環計算，將「前一狀態 (h_{t-1})」與「當前輸入 (x_t)」結合來預測「當前狀態 (h_t)」與「當前輸出 (o_t)」(如下圖)。這種循環計算的方式使得 RNN 能夠捕捉到時間序列數據中的長期依賴關係。²



RNN 可以用於各種應用領域，如文本分析、語音識別、影像識別、時間序列分析等。現在我們來看一個 RNN 應用的例子：假設我們想建構一個股票預測的模型，能利用過去 30 天的股價資料來預測明天的股價。這時我們可以使用循環神經網路 (RNN) 來實現。

首先，我們先將過去 30 天的股價資料輸入到 RNN 中，每一天的股價資料都是一個輸入。在 RNN 中，每一個輸入都會經過一個循環計算，將過去的狀態與當前的輸入結合，並且產生一個隱藏狀態。這個隱藏狀態就是 RNN 用來記錄過去輸入的資訊，並且將這些資訊用於預測當前的輸出。

在最後一天的輸入循環計算完成後，RNN 會產生一個最終的隱藏狀態，這個狀態就包含了過去 30 天所有股價資料的資訊。我們可以利用這個狀態來預測明天的股價。

雖然 RNN 有記憶能力，但實際上無法完全記得過去輸入的資訊，因此當同樣擅長處理序列的 Transformer 網路出現後，RNN 逐漸被取代。

² 具體的計算方式是：

$$\sigma(U \cdot x_t + V \cdot h_{t-1}) = h_t$$

σ 為 tanh 或 ReLU 函數

$\text{softmax}(W \cdot h_t) = o_t$ ，softmax 是一種將輸入向量正規化為每個維度總和為 1 的函數。



3.5 基於編碼器的神經網路 (BERT)

BERT (從 Transformer 生成的雙向編碼器表示技術) 是一種基於 Transformer 編碼器架構的神經網路架構，並增加額外的層和修改，使其更適合用於自然語言理解任務。BERT 在大型未標註文本資料集上進行訓練，然後可以微調用於特定任務，如文本分類與問答。BERT 使用雙向注意力，這種注意力允許模型在處理輸入時能夠考慮給定單詞左右的上下文，這比傳統的只考慮左邊上文的模型更為強大，因為單詞的意義可能取決於它左右的文本。這對於自然語言理解等任務特別有幫助，因此，BERT 在許多自然語言處理任務上取得了最優異、先進的成果。

BERT 有四種主要使用方式，可以根據目標任務選擇哪種最適合。

✓ 預訓練

BERT 可以在大型未標註文本資料集上進行預訓練，然後微調用於特定任務。

✓ 特徵擷取

BERT 可以從輸入文本中擷取特徵，這些特徵可以被當作另一個獨立模型的輸入來用。

✓ 微調

可在預訓練好的 BERT 上，新增特定任務層，在微調後用於特定任務。

✓ 轉移學習

預先訓練好的 BERT 可以作為新任務上訓練的起點，使 AI 能夠順利使用從預訓練中學到的知識。

總括來說，BERT 已被證明是自然語言處理領域裡重要的神經網路架構之一，它能夠提供有效的文本表達向量，並改善許多任務的性能。然而，要特別注意的是，BERT 需要大量的計算資源和文本才能進行預訓練，所以在某些情況下，比如處理低資源語言，可能無法使用 BERT。此外，BERT 並不是適用於所有任務的解決方案，有時可能會需要使用其他神經網路架構。

BERT 的預訓練方法

BERT 是一個廣泛使用的自然語言處理 (NLP) 模型，由 Google 開發。BERT 的一個重要特點是使用大量未經標註的文本數據進行預訓練，以學習一般語言的表徵。

在預訓練階段，BERT 被訓練來預測句中的單字。這項任務稱為「遮蔽語言建模」(Masked Language Modeling)。舉例來說，在預訓練過程中，我們會把正常的句子遮蔽一些位置，如下句：

The quick [MASK] fox jumps over the lazy [MASK] .

BERT 必須猜測被遮蔽，也就是 [Mask] 的位置是哪個單字，並根據猜測的「字機率分布」與真實的「字機率分布」去計算出兩者的損失值，用以反饋修正 BERT 的參數。以這種方式學習，BERT 就能學得語句、單字之間的語境關係，並建立對語言結構和意義的理解。

除此之外，我們通常還會採用另一個名為「下一句預測」(Next Sentence Prediction) 的任務來預訓練 BERT，也就是給定兩個句子，預測第二句是不是在文章中緊接著第一個句子，這是一個二元分類的問題，答案只有是或否。舉例來說，假設有以下兩個句子：

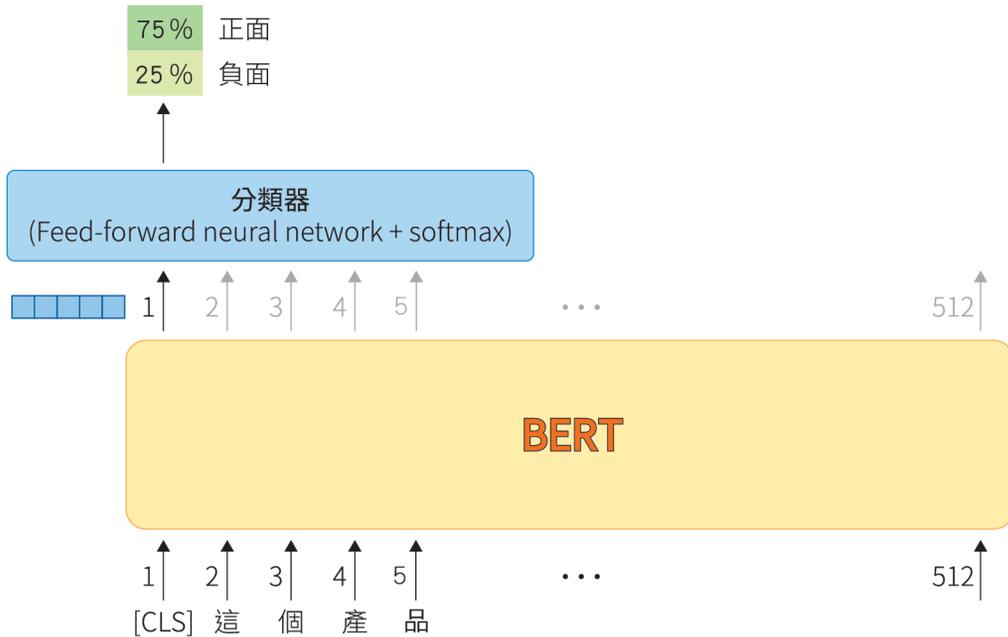
The quick brown fox jumps over the lazy dog. It barks loudly .

BERT 必須猜測第二句是不是在文章中緊接著第一個句子，並根據猜測的「分類機率分布」與真實的「分類機率分布」，計算出損失，用以反饋修正 BERT 的參數。這任務可以幫助 BERT 了解「句子之間的關係」，以及這句子如何在更長的上下文中相互配合。這對類似問答這樣的任務特別重要，因為 AI 需要了解問題和答案之間的關係，才能提供正確的回應。

綜上所述，我們在大量文本上執行「遮蔽語言建模」和「下一句預測」兩種任務，對 BERT 進行預訓練，讓 BERT 學習豐富且具有語境的語言表達向量。這些向量可以精細調整用於各種 NLP 任務，如問答和情感分析等等。

BERT 能執行的下游任務類型

微調 BERT 指的是將預先訓練好的 BERT 調整適應特定下游任務的過程。方法是在預先訓練好的網路上方增加「特定任務層」，並在「特定任務資料集」上再訓練(如下圖)。例如，當我們在 BERT 之上增加了一層分類器，再將 BERT 對整句的代表 [CLS] 特殊符號編碼所得的向量輸入，則可進行評論分類這個下游任務。



BERT 可以被微調以執行不同類型的下游任務，常見的四個類型如下：

✓ 單輸入分類

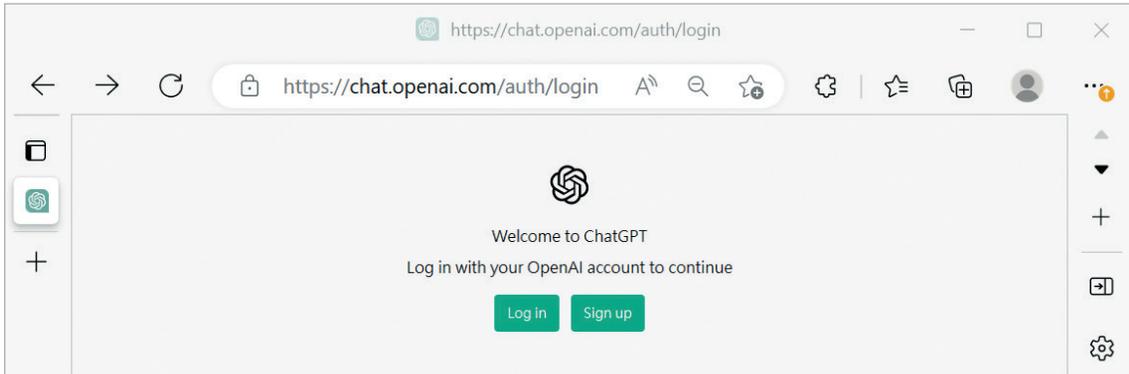
給 AI 單個輸入，再要求 AI 將其分類至已事先定義的定義類別之一。例如，我們有一批產品的客戶評論資料，希望將每筆評論分類為正面或負面。以下是基於 BERT 的 AI 執行這個任務的範例：

- 1 首先執行預處理，也就是在評論的開頭和結尾分別加上 [CLS] 與 [SEP]。
- 2 接著，BERT 將輸入文字轉換為向量序列，傳入 Transformer 層。並建立每個字與其他字之間的關係，從而理解整個句子的意義。
- 3 在 Transformer 層後，新增含有兩個輸出單元的全連接層分類器，分別對應到正面與負面。在訓練期間，分類器由標註評論資料集中，學習預測每筆評論的正確標籤（正面或負面）。
- 4 若要對新的評論進行分類，可以透過相同的預處理步驟，將資料傳入 Transformer 層，可得到 [CLS] 編碼所得的向量，再輸入分類器，就可以得到正面與負面的機率。

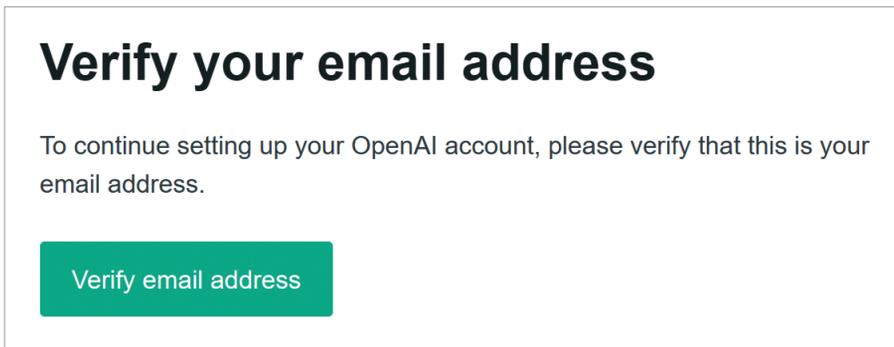


操作體驗

- 1 首先，在 Google 輸入 chat openai，點擊排名第一的網址，接著應該會進入到以下畫面：



- 2 若有 OpenAI 帳戶，請點擊左邊的「Log in」，就可以進入 ChatGPT 網站；沒有 OpenAI 帳戶，請點擊右邊的「Sign up」，進行註冊動作。
- 3 請輸入你的 email 帳號，並點擊「Continue」。
- 4 請輸入你預備使用的密碼，並點擊「Continue」，OpenAI 系統會寄一封請你確認的 email 給你（如下圖），請點擊「Verify email address」。



- 5 點擊以後，瀏覽器會跳轉進入輸入姓名的畫面。當你完成輸入姓名後，請點擊「Continue」，緊接著就會跳轉進入輸入電話號碼的畫面。當你完成輸入電話號碼後，請點擊「Continue」，系統就會要求你輸入 code。請到手機上查看簡訊，看看 code 是多少，當你完成輸入 code 後，就會進到 ChatGPT 網站了。



DeepLab¹⁸ 是由 Google 開發的一系列圖像分割網路。它使用具有編碼器—解碼器結構和擴張卷積 (Dilated Convolution) 的卷積神經網路，來學習輸入數據的豐碩特徵。DeepLab 在多種圖像分割評測上取得了最優質的成果。

以上每個網路都是基於 CNN 神經網路的架構發展出來的深度神經網路，由於架構都較為複雜，這裡僅作概略性的說明，目的是為了讓讀者知道電腦視覺領域中三種主要任務表現較出色的系列網路。若讀者有興趣，可以深入閱讀原始論文以得到最精確的資訊。

3.8 生成網路

生成對抗網路 (Generative Adversarial Networks, GANs)

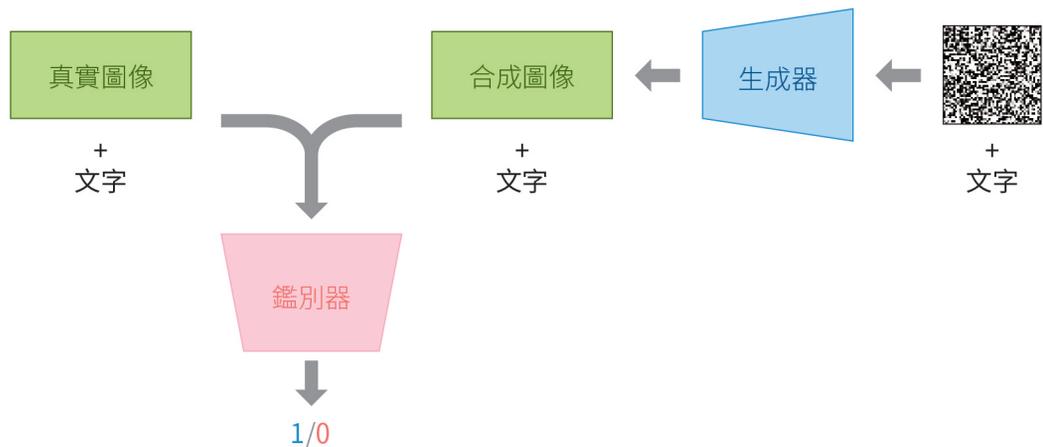
生成對抗網路 (GANs) 是 2014 年由 Ian Goodfellow 等人提出的神經網路架構¹⁹。其中包含兩個單獨的模型：生成器和鑑別器。生成器的目標是生成與訓練資料集裡類似的合成數據，而鑑別器的目標是區分生成器生成的合成數據和訓練資料集中的真實數據。

以文字到圖像的 GAN 為例，其目標是生成與給定文本描述的相關圖像。訓練資料集包含文字描述和對應圖像的配對，生成器負責把給定的文字描述生成圖像，而鑑別器負責確定生成的圖像是否為真實。

生成器和鑑別器是兩個相互競爭的網路，在敵對過程中同時訓練。生成器會試圖生成逼真到能愚弄鑑別器的合成圖像，而鑑別器則試圖將生成器合成的圖像正確分類為假的。這個過程持續進行，直到生成器能夠生成與真實圖像無法區分的高品質圖像，此時 GAN 就會被認定已訓練完成。

18 Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4), 834-848.

19 Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.



整個過程如上圖所示：

- 1 生成器輸入文字描述並生成合成圖像。
- 2 鑑別器輸入 [真實圖像 + 文字] 與 [合成圖像 + 文字]，並試圖確定兩者之間的區別。
- 3 生成器根據其成功欺騙鑑別器的程度去接收反饋並更新其參數。
- 4 鑑別器根據其成功區分合成圖像和真實圖像的程度去接收反饋並更新其參數。

重複執行這個過程，直到生成器能夠生成與真實圖像無法區別的高品質圖像。然後，最終得到的 GAN 就可以用於在給定文字描述的情況下生成新圖像。

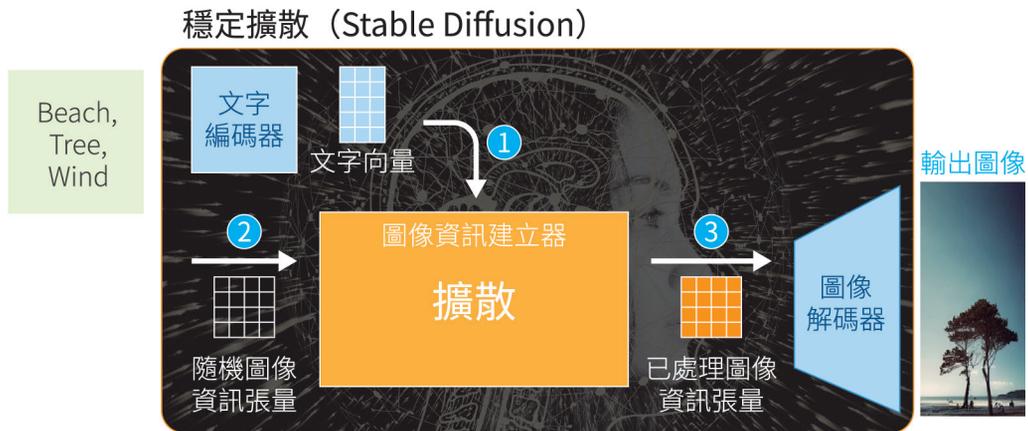
為了讓讀者能較快了解文字到圖像的 GAN，上述的過程略過了某些細節。如果讀者想知道完整的過程，可以參考原論文。²⁰

²⁰ Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016, June). Generative adversarial text to image synthesis. In *International conference on machine learning* (pp. 1060-1069). PMLR.



穩定擴散網路

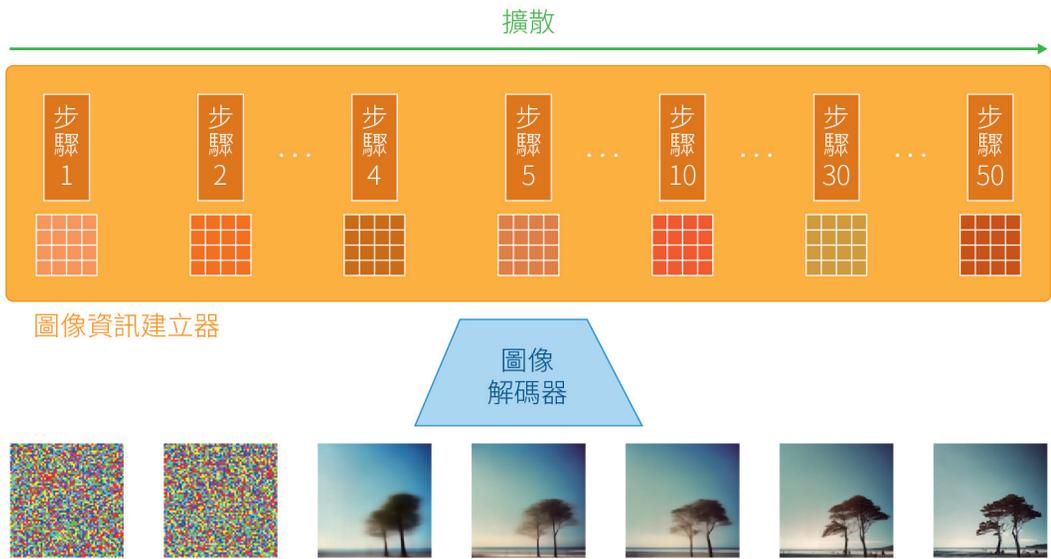
穩定擴散網路 (Stable Diffusion) 於 2022 年於 CVPR 研討會中發表²¹，是一種接受提示文字，進而產生圖像的生成網路 (如上圖)。這個網路由「文字編碼器」、「圖像資訊建立器」、以及「圖像解碼器」所組成。每當接收到提示文字，文字編碼器就會將提示文字轉為「文字張量」，也就是一組向量，與此同時，網路內部也會隨機產生一個「圖像資訊張量」；接著，「文字張量」(1) 與「隨機圖像資訊張量」(2) 一起輸入「圖像資訊建立器」，在建立器中進行擴散，最後產生「已處理圖像資訊張量」(3)，由「圖像解碼器」解碼成為輸出的真實圖像。



以下分別敘述這三個元件的功能細節：

- ✓ 文字編碼器用來將輸入的提示文字轉為張量，也就是一組向量。
- ✓ 圖像資訊建立器經過 50 個步驟後，在「圖像資訊空間」中漸進式地將初始的「隨機圖像資訊張量」轉變為最終的「已處理圖像資訊張量」(如下圖)。如果我們將每一個步驟的「圖像資訊張量」透過「圖像解碼器」來解碼，就可以觀察到每一個步驟的圖像資訊張量所對應到的真實圖像，且雜訊也在經過 50 個步驟後，逐漸減少。

21 Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10684-10695).



在每個步驟中，前一個步驟的「圖像資訊張量」、「文字張量」以及「步驟值」會一起輸入雜訊預測器中，產生現在這一個步驟的圖像資訊張量（如下圖）。例如，當我們要描繪出步驟3的情況，可以看到步驟2的圖像資訊張量、文字張量、以及步驟值3一起輸入「雜訊預測器」中，緊接著就產生步驟3的圖像資訊張量。



- ✔ 圖像解碼器負責將「圖像資訊張量」解碼為真實圖像。

透過上述的整個過程，Stable Diffusion 就能夠將輸入的提示文字轉成一張圖像。限於本書的定位，這裡僅介紹由提示文字到產生圖像的過程，建議想深入研究的讀者可以參考原始論文。



應用：AI 算圖服務 DALL·E 2

Open AI DALL·E 2 是一個最近推出的人工智慧技術，它可以通過圖像生成器來創建獨特且極具創意的圖像。DALL·E 2 使用深度學習技術，並且可以理解自然語言的描述，通過將這些描述轉換為圖像，創造出令人驚嘆的視覺效果。不僅如此，DALL·E 2 還可以通過修改圖像的不同屬性，如顏色、大小和形狀等來生成各種變化。這項技術將為人們提供更多創意和美學的可能性，並將應用在多個領域，如設計、藝術和電影等。

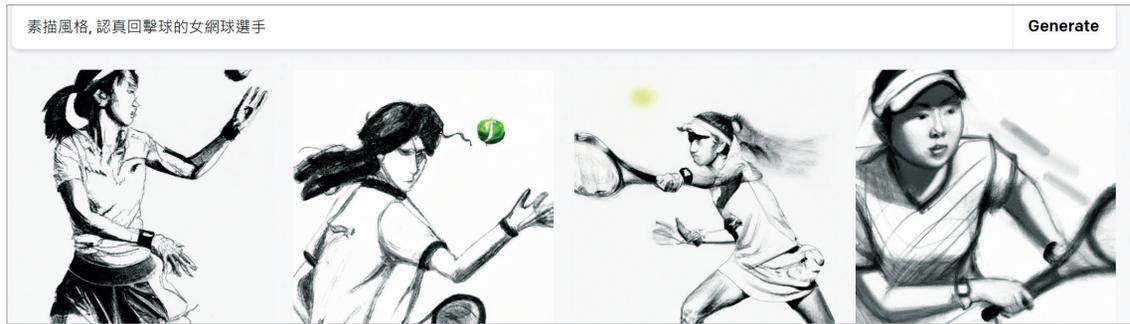
登錄 DALL·E 2 網站的步驟如下：

- 1 首先，在 Google 搜尋 DALL E 2，點擊傳回的第一個網址。
- 2 點擊網頁右上角的 "Sign In" 按鈕，進入登錄頁面。
- 3 如果您還沒有 DALL·E 2 的帳戶，可以點擊 "Sign Up" 按鈕，填寫相關的註冊資料，建立新的帳戶。如果您已經有帳戶，請輸入您的電子郵件地址和密碼，點擊 "Sign In" 進行登錄。
- 4 如果您使用 Google 等第三方帳戶，也可以點擊 "Sign in with Google" 按鈕進行登錄。成功登錄後，您可以開始使用 DALL·E 2 生成器進行創作。

下面是操作 DALL·E 2 網站生成圖像的步驟，以創建一張植物造型為例：

- 1 登錄 DALL·E 2 網站後，在文字輸入框中，輸入想要創作的圖像的描述，例如 "一棵帶葉子的綠色植物"。
- 2 點擊輸入框右方的 "Generate" 按鈕，Dall-e2 會根據您的描述生成一張圖像如下圖。假設最喜歡第四張，點及第四張圖。
- 3 點擊右上角的  按鈕，將圖像下載。
- 4 您可以在生成器頁面右方查看所有曾經生成的圖像，並下載、分享或刪除它們。

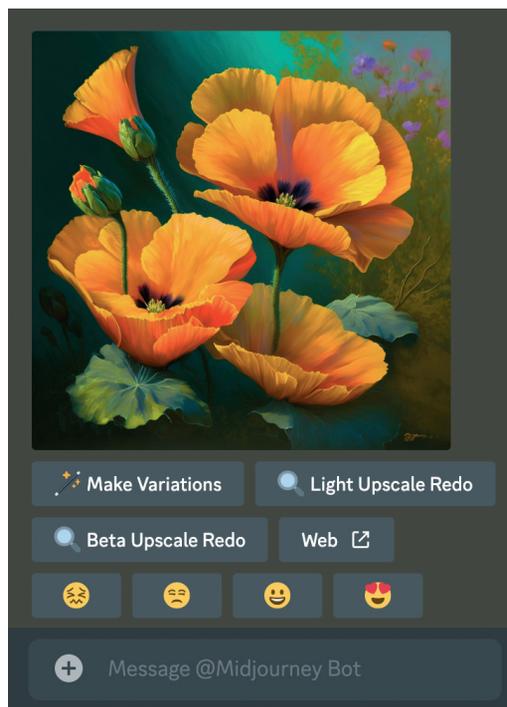
以上就是使用 DALL·E 2 網站生成圖像的基本步驟，透過這樣簡單的操作，您就可以創造出各種獨一無二的圖像。



Midjourney

Midjourney 是一個獨立的研究實驗室，致力於探索新的思想媒介和拓展人類的想像力。據傳，Midjourney 早期的版本曾經採用 GAN 技術訓練算圖模型，而最新的版本則是採用穩定擴散架構。

進入 Midjourney 後，只要輸入「/imagine」，就可以開始輸入給 Midjourney 的提示 (prompt)，提示的內容必須用白描、樸素、簡練的寫法敘述出您腦中的景象。接著按下「送出」，Midjourney 就可以為您產生一張圖像 (如下圖)。如果要更專業些，還可以在後面加上更多風格、解析度、光線等等的修飾字，不論是照片風、寫實風、日本動漫風、好萊塢 3D 動畫風、超現實風，都可以為您產生出來。





詳細用法，請參見 Midjourney 官方網站²²或加入 Midjourney 社群，可以得到更豐富的資訊。

3.9 多模態深度學習

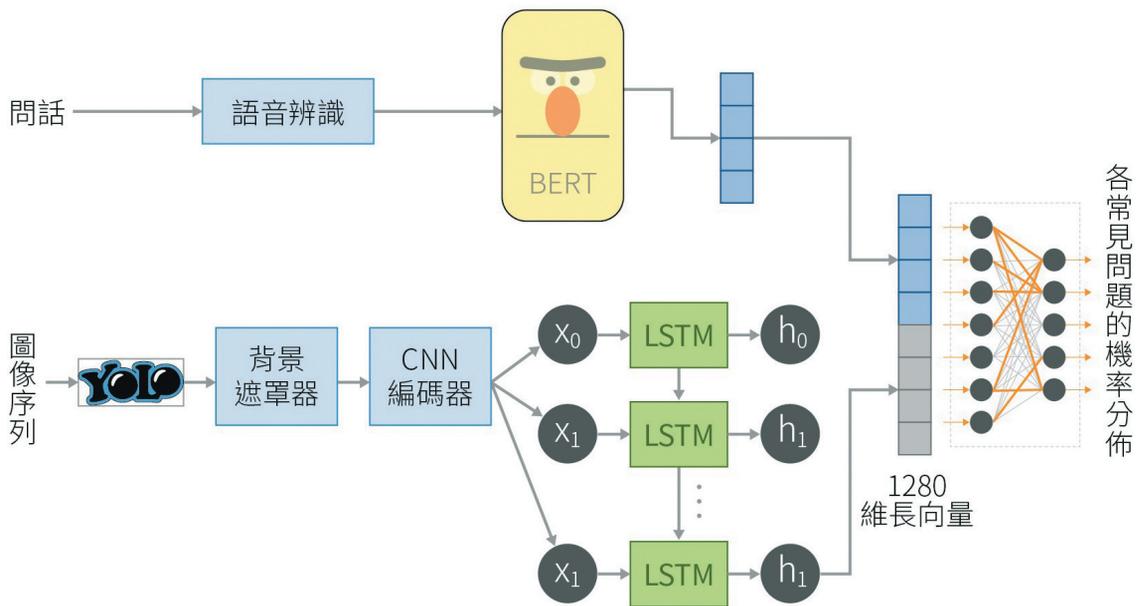
多模態深度學習 (Multimodal deep learning) 是指使用能夠處理和整合多種類型的資料 (例如文字、圖像和聲音等等) 的深度學習。近年來，這種學習方式越來越重要，因為許多真實世界的問題皆涉及多種模態的資料。

多模態深度學習常常應用在電腦視覺領域。例如，我們可以設計一個 AI，它能根據外觀與伴隨的文字描述來分類動物圖像。這個 AI 可以被訓練來辨識圖像中的特定特徵，譬如皮毛圖案或面部特徵，並使用文字描述來消除相似外觀的動物彼此之間的歧異。另一個常見的應用是在自然語言處理領域。我們可以設計一個 AI，能依據語音與隨同的文字來即時生成精確的翻譯。

總之，因為多模態深度學習能夠利用多個資訊來源，所以提高了深度學習模型在各種任務上的表現。

接著我們以蔡宗翰教授 (AI 界李白老師)、李宏毅教授、畢南怡博士等人於臺灣大學 IoX 創新研究中心合作發展的多模態組裝助教對話機器人，來說明如何應用多模態深度學習。這個組裝助教對話機器人能即時分析由視訊鏡頭得到的短片，判斷組裝者想詢問的問題。其架構如下：

22 <https://docs.midjourney.com/>



假設學生在組裝時發生疑惑（如下圖（a）），並問了以下問題（如下圖（b））：

Is it the right direction to lock the screw?

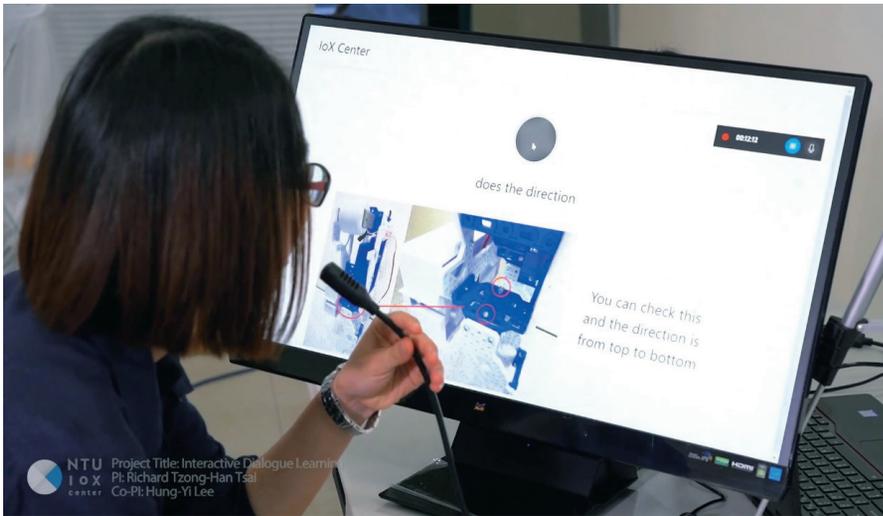
學生的問話透過語音辨識之後轉換成文字，再輸入到目前最強大的語言模型 BERT 中轉換成代表語言的向量。另一方面，學生問話的短片也會轉成連續的圖像，每張圖像都會經過 YOLO 模型辨認組件，並將背景去除，最後透過 CNN 編碼器轉為向量。於是，這個短片就會被轉換成一系列的向量，再輸入 RNN 網路中，得到代表影像的向量。最後，將語言及圖像的向量接合起來，成為 1280 維的長向量，輸入一個全連接神經網路進行分類後，就可以得到該問話屬於各常見問題的機率。系統就可以將具有最高機率的常見問題對應到的答案，以語音、文字、圖像的方式合併呈現給學生（如下圖（c）），學生就可以根據機器人的回應，繼續進行組裝工作。若讀者對這個專案的細節感興趣，可以閱讀本專案發表在 AAI 的短論文²³與 IEEE 的長論文²⁴。

²³ Chiu, Y. C., Chang, B. H., Chen, T. Y., Yang, C. F., Bi, N., Tsai, R. T. H., ... & Hsu, J. Y. J. (2021, May). Multi-modal User Intent Classification Under the Scenario of Smart Factory (Student Abstract). In *Proceedings of the AAI Conference on Artificial Intelligence* (Vol. 35, No. 18, pp. 15771-15772).

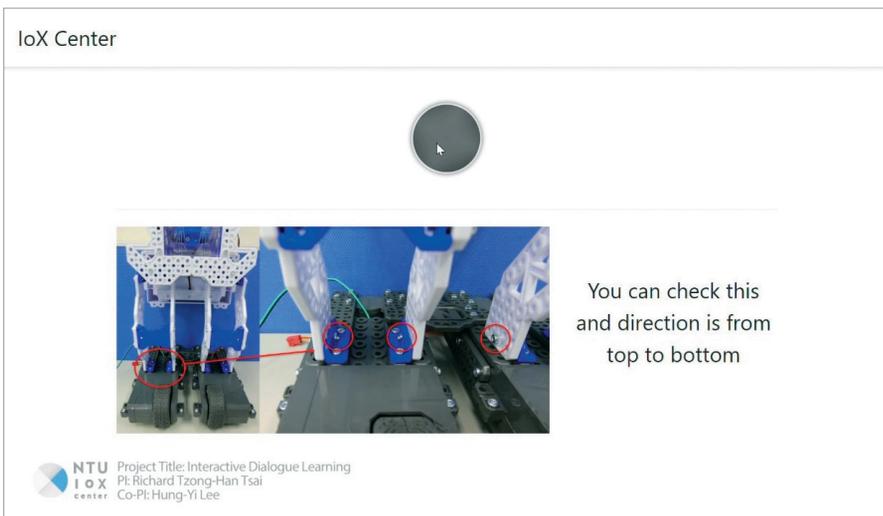
²⁴ Chen, T. Y., Chiu, Y. C., Bi, N., & Tsai, R. T. H. (2021). Multi-modal chatbot in intelligent manufacturing. *IEEE Access*, 9, 82118-82129.



(a) 在組裝時，組裝者對上方部件的方向產生疑惑。



(b) 組裝者詢問組裝助教對話機器人。



(c) 裝助教對話機器人以語音指導學生，使學生能夠參照畫面所指示的方向組裝上方部件。

INTRODUCTION TO
AI+AIOT



PART

II

AI 的各種應用

第 4 章 人文與法律應用

第 5 章 醫療應用

第 6 章 金融應用

第 7 章 行銷零售應用

第 8 章 工業應用

第 9 章 農業應用

第 10 章 未來展望

這一部分的第四章到第十章談到 AI 的各種應用，包含 AI 結合物聯網收集數據之後的 AIoT 應用。這部分做法是針對人文與法律、醫療、金融、行銷與零售、工業，以及農業應用展開，最後以科技趨勢思考為最後章節，討論永續、元宇宙，以及產業 AI 化進階應用。說明如下：

✔ **第四章：人文與法律應用**

這章實務上的應用較少，以蔡宗翰老師團隊完成的案例做法為案例說明。

✔ **第五章：醫療應用**

這章包含了醫療紀錄分析、加速藥物開發、醫療影像分析，以及風險預警與其他應用的說明。而醫療紀錄分析更有蔡宗翰老師團隊完成的案例協助說明。

✔ **第六章：金融應用**

這章包含了提升客戶體驗、信用與貸款決策、AI 軟硬體機器人應用、法遵與風險控管，以及精準行銷等應用的說明。

✔ **第七章：行銷與零售應用**

這章包含了全通路智慧零售、無人機與機器人應用，以及行銷推廣等應用的說明。

✔ **第八章：工業應用**

這章包含了電腦視覺輔助作業、生產大數據建模提高良率、工業機器人的應用，以及數位孿生與設備預測性維護等應用的說明。

✔ **第九章：農業應用**

這章包含了精準農業、畜牧業與養殖漁業的人工智慧、農漁牧業機器人解決方案，以及智慧營運與運輸等應用的說明。

✔ **第十章：未來展望**

這章談到了永續發展（ESG）、元宇宙，以及產業 AI 化進階應用的發展可能等的相關說明。

大家可以針對自己有興趣的部分開始閱讀吧！

06 金融應用

6.1 介紹

智慧金融是因應全球數位化趨勢而針對金融業應用的科技，人工智慧在其中扮演很重要的角色：透過人工智慧的技術，提供客戶更好、更符合需求的金融產品。而如何讓客戶有好的體驗，讓客戶黏著度高，甚至推薦朋友成為金融機構的新客戶，是很重要的事。

透過攝影機在 VIP 客戶到機構場域中時，第一時間由攝影機搭配影像辨識，立刻通知相關人員 VIP 客戶到場，接下來可以提供此客戶親切的服務，讓客戶有好的感受，相對關係與生意就比較能長久！

另外很多在智慧型手機的支付服務及 ATM/VTM 上的金融服務，需要更精準的人體特徵辨識，這也是人工智慧與物聯網可以著力的地方。

客戶在網頁上的交易軌跡、電話中的服務錄音、客戶的實體交易紀錄…等等數據，結合金融機構跟客戶有關的其他數據，做人工智慧分析，就可以達成更瞭解客戶，因而針對客戶做到精準行銷。

另外，人工智慧可以增加營運效率，達成自動化，甚至強化對各種風險的防範。



6.2 AI 應用架構與案例

為了提升客戶體驗，金融業大量使用人工智慧：利用物聯網裝置收集到客戶的接觸資料，分析以瞭解客戶行為，作為客戶視圖、客戶接觸據點設立依據；而具備人工智慧的無人機更可以作為客戶理賠時第一時間到達客戶災損所在勘查的重要工具。金融業也開始導入 Chatbot 及理財機器人兩種人工智慧軟體機器人，接下來使用的企業會越來越多。另外人工智慧還有用各種數據資料作出信用模型以協助貸款、對客戶精準行銷產品，以及協助法遵、資訊安全防禦與風險管控。

因為智慧金融包含的除了有提升客戶體驗，信用與貸款決策、軟硬體機器人應用（包含無人機）、法遵與風險控管，以及精準行銷，所以本章規劃如圖 6.1。

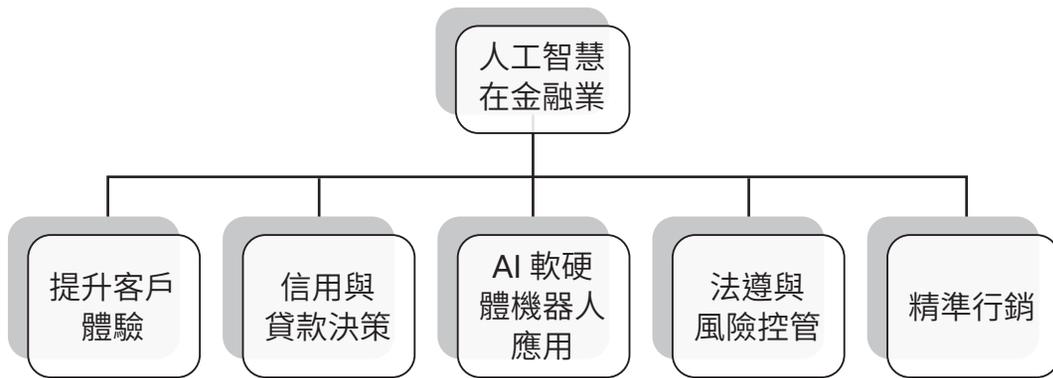


圖 6.1 智慧金融的人工智慧應用範疇

圖片來源：裴有恆製

6.2.1 提升客戶體驗

網際網路顛覆了消費習慣，消費者開始不必去銀行，隨時隨地都可以上網查資訊、轉帳與買進股票及金融資產；保險不只是在網路上可以購買意外險等簡易險，營業員在疫情期間利用線上視訊跟客戶對談買保險，就算是面對面討論，都會利用平板連網提供客戶資訊及紀錄作業。聯網相關的方便性與易於管理，讓越來越多族群在網路上進行金融交易與處理。透過網路與人工智慧的，實體與虛擬融合的金融科技已經是新的趨勢了，所以真正的商機會在如何提升消費者的體驗，讓消費者願意做更多金融交易，購買更多金融商品。

客戶體驗的強化跟客戶隱私的保護如何兼顧，一直是金融業的兩難，因為體驗要強化，要有相關的隱私權的揭露，生物特徵相關之辨識技術已有導入客戶服務之實務做法，考量生物特徵為客戶隱私之一部分，在前任金管會主委顧立雄的任職期間，讓法規鬆綁為 106 年 01 月 10 日發表的「為強化金融服務生物識別運用之相關安全控管」法規：「銀行業者在運用客戶生物識別資料時，內部作業及資料之保存應有嚴謹之控管程序；取得及利用客戶生物特徵資料前，應先取得客戶同意並留存客戶同意之紀錄，以避免爭議，俾符合個人資料保護法之規定。」

中國大陸因為對個人隱私比較沒有像台灣這麼嚴的管制，所以相關的金融科技導入也較早。這裡我們舉了中國大陸天誠盛業銀行以人臉識別強化客戶服務、螞蟻科技集團的支付寶的刷臉支付服務。

案例（一）：天誠盛業銀行的 VIP 客戶人臉識別解決方案

中國大陸的天誠盛業銀行的 VIP 客戶人臉識別解決方案，基於其研發的 SmartBIOS，以人臉識別技術 TesoFace V4.0 為核心。

系統運作流程為透過攝影機去讀取客戶的人臉，與預先保存的 VIP 客戶人臉模板比對，在第一時間識別出到訪的 VIP 客戶。然後系統會透過語音方式歡迎 VIP 客戶的光臨，結合後台發送的簡訊、手持裝置訊息推送等方式，即時將 VIP 客戶到達訊息發送給對應的客戶經理好出門迎接，使 VIP 客戶深切感受到“賓至如歸”的優越服務，提升 VIP 客戶體驗。

這套系統目前已經使用在中國大陸民生銀行。¹

案例（二）：螞蟻科技集團的支付寶的刷臉支付 Smile to Pay

螞蟻科技集團以支付寶起家，2014 年，阿里巴巴集團分拆旗下金融業務，成立浙江螞蟻小微金融服務集團股份有限公司（簡稱螞蟻金服），2020 年 6 月變更為螞蟻科技集團股份有限公司。²

1 資料來源：天誠盛業官網。

2 資料來源：Wikipedia <https://zh.wikipedia.org/zh-tw/%E8%9A%82%E8%9A%81%E9%9B%86%E5%9B%A2>



針對生理資訊認證，支付寶已經支援客戶臉部辨識支付（刷臉支付），這是應用人工智慧的機器學習能力。

2017 年，連鎖餐飲集團百勝中國與螞蟻金服合作，在 KFC 子品牌全新餐廳概念店 KPRO 推出刷臉支付服務，這是全球首次應用刷臉支付 Smile to Pay，免用手機，透過臉部辨識就可完成支付動作。

馬雲在 2015 年 3 月的德國漢諾威消費電子展（CES）上，就展示了支付寶的刷臉支付技術，而經過兩年多的小心測試，並且結合 3D 攝影機來強化支付技術，才真正實際商用。

透過 3D 成像，3D 攝影機能夠辨識視野內空間每個點的三維座標，讓電腦得到空間的 3D 數據，並能夠重建完整的三維世界，實現三維定位，也因此人臉辨識功能可以分辨出平面圖像 / 平面影像 / 化妝 / 皮面具 / 雙胞胎等複雜狀態。

現在這樣的服務在很多阿里巴巴相關企業導入。



圖 6.2 顧客在 KPRO 餐廳使用支付寶 Smile to Pay 刷臉支付功能支付餐費

圖片來源：<http://www.yumchina.com/News/Index/2017-1-255>

6.2.2 信用與貸款決策

而因為有了交易的紀錄，就可以根據買賣交易紀錄做信用評等及相關模型，來根據對應的模型來做貸款決策。台灣的財團法人金融聯合徵信中心（簡稱聯徵中心）的個人信用評分是由會員金融機構定期報送有關個人的最新信用資料，依受評分的對象的特性，套用對應適用的信用評分模型，而根據聯徵中心網站上的資料得知「個人信用評分模型採用的資料，大致可區分為下列三大類：

✔ 繳款行為類信用資料

係指個人過去在信用卡、授信借貸以及票據的還款行為表現，目的在於瞭解個人過去有無不良繳款紀錄及其授信貸款或信用卡的還款情形，主要包括其延遲還款的嚴重程度、發生頻率及發生延遲繳款的時間點等資料。

✔ 負債類信用資料

係指個人信用的擴張程度，主要包括負債總額（如：信用卡額度使用率，即應繳金額加上未到期金額 ÷ 信用卡額度；如：授信借款往來金融機構家數）、負債型態（如：信用卡有無預借現金、有無使用循環信用；如：授信有無擔保品）及負債變動幅度（如：授信餘額連續減少月份數）等三個面向的資料。

✔ 其他類信用資料

主要包括新信用申請類之相關資料（如：金融機構至聯徵中心之新業務查詢次數）、信用長度類之相關資料（如：目前有效信用卡正卡中使用最久之月份數）及保證人資訊類相關資料等。³

因為以聯徵中心的資料為主要資料，而剛出社會沒有信用資料的人很難套用這個模型，所以需要導入其他方式協助。在 AI 時代，因為可以收到的數據很多，特別是電商與社群數據，可以協助原有的模型強化。

這裡以螞蟻金服推出芝麻信用，以及數位融資平台 Funding Societies 提供的服務為例，說明他們如何用各種數據來建立模型。

3 資料來源：聯徵中心官網

案例（四）：數位融資平台 Funding Societies

總部位於新加坡的 Funding Societies 是由 Kelvin Teo 和 Reynold Wijaya 共同創立於 2015 年 2 月。⁵ Funding Societies 提供從 500 美元到 150 萬美元不等的貸款金額，而其客戶有微型的社區型商店、電商供應端、正經歷快速增長階段的新創公司、不想耗時兩到三個月才能獲得銀行貸款的老公司等。

Funding Societies 選擇利用交易訊息、線上評論和供應鏈數據等數據，來訓練信用評分模型；接著，再透過多元的業務覆蓋率，取得客戶購置成本 (Customer Acquisition Cost, CAC) 和貸款價值比 (Loan-to-value ratio, LTV)。

雖然 Funding Societies 的利率普遍高於銀行，但低於或等於信用卡，也提供附有簽帳金融卡功能的信用卡來替代公司卡，並與蝦皮、印尼電商 Bukalapak、簿記應用程序 BukuWarung、金融科技 Alterra 和農業科技平台 Tanihub 進行企業合作，提供營運資金貸款給中小企業客戶。除了持續蒐集的貸款償還數據資料外，公司的部分數據是透過合作關係取得專屬使用權。

Funding Societies 的主要市場為印尼，但在新加坡、馬來西亞和泰國都已獲得註冊和執照許可，也展開在越南的營運，並打算進入菲律賓。自推出以來，已通過超過 490 萬筆貸款，並向中小企業發放總計超過 20 億美元的貸款。⁶

6.2.3 AI 軟硬體機器人應用

有了人工智慧，當然就會運用人工智慧的強大運算能力，進行快速的處理，快速的自動化交易這時就是必然做的項目。玉山銀行前科技長陳昇瑋就在 Fintech Taipei 2019 研討會講題「論 AI 與機器學習的特徵與應用現況」中提到「分析證券市場時，交易員必須學習基本面分析、技術面分析、特徵工程等技術與交易規則；但規則對 AI 來說並不是必需的，只需要提供財報、籌碼等市場資訊，AI 就可以自己做預測，自己找出金融市場的交易規則。而 AI 用財經新聞做預測隔日股價漲跌的話，準確度可達 80% 到 85%。目前全球股市已經有 8 到 9 成交易是程式執行的，而全球十大對沖基金中，有 6 個導入 AI 協助交易。除了證券交易，AI 也被用在產業分析，用來預測股價。」⁷ 理財機器人就是這樣的機制，這裡以富邦的理財機器人為例。

5 資料來源：Wikipedia https://en.wikipedia.org/wiki/Funding_Societies

6 資料來源：數位時代 <https://www.bnext.com.tw/article/67925/funding-societies-asco>

7 資料來源：TechOrange <https://buzzorange.com/techorange/2019/11/30/fintech-taipei-ai-in-finance/>



使用無人機與機器人減少人力需求，增加客戶體驗，已經是現在進行式，未來更因為科技進步，無人機與機器人將會更靈活、聰明，可以看到這樣發展的潛力。這樣的機器人有實體的，如服務機器人 Pepper 或無人機，也有虛擬的，如做客服的玉山銀行的隨身金融顧問。而在 2022 年底開始測試，2023 年紅透半邊天的 ChatGPT 等生成式 AI 模型因為反應很人性化，內容又完整，可以創造更好的客服體驗。

另外，透過機器人流程自動化 RPA⁸，因為協助相關輸入、輸出行政流程，也是在金融業及各行各業盛行的機制。

案例（五）：台新銀行 Richart 數位帳戶與智能投資

Richart 數位帳戶是台新銀行於 2016 年推出的數位帳戶，主打金融結合生活，透過網路，就能滿足大多數的金融需求，可以說是把實體銀行搬到網路上，方便是它的主要特色。與台新銀行一般帳戶主要差別在於，沒有實體存摺，所以沒辦法帶著存摺到銀行領錢，但一樣有提款卡，可以到自動櫃員機（ATM）提款或存款。⁹

Richart 到本書完稿前是台灣 No1 的數位帳戶，因為其把所有的服務都整合在小小的 APP 之中，所以任何你想到的服務都可以被滿足，而其透過人工智慧的智能投資，只要 10 元或 1 美元就可以開始投資基金，自動推薦投資組合，也可以進行購買 ETF。¹⁰

Richart 的智能投資是照著「分析」、「監控」、「調整」三步驟，並透過人工智慧學習，進而預估未來市場走勢，可迅速獲得合適的基金組合。而其標的橫跨全球熱門商品，提供包括新台幣、美元雙幣別投資股票型、固定收益型等基金類型商品。只要在事前設定好停利停損條件，當發生投資人所設定的觸動條件情境，Richart 系統就會自動在智慧型手機上彈出調整投資組合的提醒，幫投資人做財富把關。¹¹

8 英文 Robotic process automation，簡稱 RPA，是以軟體機器人及人工智慧為基礎的業務流程自動化科技。資料來源 Wikipedia <https://zh.wikipedia.org/zh-tw/%E6%A9%9F%E5%99%A8%E4%BA%BA%E6%B5%81%E7%A8%8B%E8%87%AA%E5%8B%95%E5%8C%96>

9 資料來源：麻布記帳網頁 <https://blog.moneybook.com.tw/?p=1>

10 資料來源：這就是人生 <https://www.beurlife.com/2021/12/taishin-richart-digital-bank.html>

11 資料來源：Yahoo! 股市 <https://tw.stock.yahoo.com/news/%E8%A1%8C%E5%BA%AB%E5%8B%95%E6%85%8B-%E5%8F%B0%E6%96%B0richart%E6%8E%A8%E5%87%BA-ai%E6%99%BA%E8%83%BD%E6%8A%95%E8%B3%87-%E4%B8%8A%E7%B7%9A-%E5%80%8B%E6%9C%88%E5%B9%B4%E8%BC%95%E7%94%A8%E6%88%B6%E6%88%90%E9%95%B71-042508018.html>

案例（六）：富邦證券的理財機器人

富邦證券與工研院及麻省理工學院（MIT）專家團隊——訊能集思智能科技有限公司（Synergies Intelligent Systems, Inc）合作，打造「富邦理財機器人」，透過智慧工具與大數據精算，提供定期定額買台股（ETF）投資標的組合建議。富邦證券希望透過提供「富邦理財機器人」的服務，提升服務滲透率，讓高資產或者低交易頻率的客戶，都可以依據個人的資產狀況，透過智慧機器人來協助他們，達成理財需求。

這套系統整合機器學習、人工智慧、數據分析等眾多技術，只要透過富邦證券的「WEB 網路交易系統」，或是「富邦 e+」App，即可立即進入「富邦理財機器人」服務平台，以設定投資目標與投資金額。而投資人只要點選 KYC 測驗，並選擇符合需求的理財計劃，透過 KYC 測驗的問答以及您所選擇的投資目標¹²，理財機器人使用機器學習方法了解您的風險屬性、投資偏好、資金需求。系統採用發源自高盛集團的 Black-Litterman Model，不僅顧及了報酬率以及投資風險，還能考量多個投資觀點面向，克服現代投資理論對價格波動過於敏感的問題，好建構更有效率、適合長期投資的投資組合建議。¹³

案例（七）：玉山銀行的 ChatBot「玉山小 i」隨身金融顧問

玉山銀行與台灣 IBM、LINE 合作，在 2017 年 4 月 25 日推出 AI ChatBot「玉山小 i」隨身金融顧問。透過 IBM Watson Conversation 支援繁體中文的自然語意分析技術，與玉山銀行打造出基於 LINE 平台的即時互動聊天機器人服務，提供個人化金融服務建議。

「玉山小 i」將 AI 技術結合至外匯諮詢、房貸評估、信用卡推薦等金融諮詢服務，且能夠與顧客自然對話，提供用戶專屬的諮詢服務。

這項服務的背後憑藉的是 IBM Watson Conversation Service，它於 2016 年 7 月在 IBM Watson 開發者雲網站（Watson Developer Cloud）上推出，目的是協助企業端與開發者在 IBM 雲端平台上快速打造、測試、部署自己的 AI ChatBot 應用；透過蒐集中文對話內容進行語音的分析與辨識，而能產生接下來的對話內容。

12 資料來源：聯合報報導 <https://udn.com/news/story/7255/2656815>

13 資料來源：富邦證券官網



不僅如此，透過機器學習持續式的訓練擴增智慧，並加入玉山銀行行員的專業知識與對話式服務，「玉山小 i」會不斷精進與應對互動的服務品質，提升顧客滿意度。¹⁴



圖 6.3 玉山銀行的玉山小 i 隨身金融顧問使用

圖片來源：裴有恆擷圖

6.2.4 法遵與風險管控

在立法院全球資訊網中強調「法令遵循機制係內部控制制度中重要的一環。所謂法令遵循機制係指在事前將各項法令及銀行內部規章等規範進行辨識，並與銀行內部相關單位溝通、衡量法令之落實，進而提出建議，監督銀行之運作皆符合法令規範之機制。」¹⁵ 由此可知以 AI 科技做好法遵的法遵科技，因為可以幫助銀行做到繁複的法令遵循，避免因為觸法而付出高額罰款。

14 資料來源：科技報橘報導 <https://buzzorange.com/techorange/2017/04/25/ibm-line-esunbank-chatbot/>

15 資料來源：立法院官方網站 <https://www.ly.gov.tw/Pages/Detail.aspx?nodeid=6590&pid=85277>

風險管控對金融機構是很重要的部分，像是 AML (Anti-Money Laundering)，也就是反洗錢，是指金融機構和政府用於預防和打擊金融犯罪（尤其是洗錢和恐怖主義融資）的總體監控的措施和過程¹⁶。台灣在 107 年 11 月通過了洗錢防制法，在其中第一條就說明，「為防制洗錢，打擊犯罪，健全防制洗錢體系，穩定金融秩序，促進金流之透明，強化國際合作，特制定此法」。之前台灣兆豐銀行紐約分行，在 2016 年 8 月 19 日因違反反洗錢防制法，慘遭紐約州金融服務署處以台灣金融史上最高罰款 1.8 億美元震撼金融界。¹⁷而人工智慧可以大大強化這個部分，我們以玉山金控為案例跟大家說明。

資安是企業的另一大風險，台灣的金融業更是從第一銀行海外分行被駭客入侵後用 ATM 盜取銀行存款開始被矚目。金融業更是駭客入侵的幾大首要目標之一，特別是現在科技日新月異，所以這邊介紹以 AI 防護著稱的奧義智慧的 AI 資安工具。

案例（八）：永豐金控的 AI 法遵智能平台

永豐金在 2021 年 10 月 13 日宣布自家導入法遵智能平台，包括法規智能分析系統、法遵風險評估系統及法遵管理儀表板。

此法遵平台透過賦能與創新科技，將繁複的法令遵循作業智能化及自動化，除可自動蒐集包括金管會、證券暨期貨法令查詢系統等 20 餘個資料源的法令、函釋及裁罰資料，而且每日進行更新。還有導入了 AI 智能及 rule-base 演算法交叉比對技術，將數千部外部監管規範與公司內千餘部內規，建立高密度的「條對條」的關聯比對，一旦監管規範有異動，系統就會通知進行內規檢視，達到自動、即時及精準監控的效果。

此外，該平台也建置法令遵循風險評估（CRA）系統，因應法遵風險評估及管理為主管機關近年來的監理重點，透過此系統，永豐金旗下銀行及證券子公司以智能化方式，執行法令遵循風險評估作業。

16 資料來源：<https://medium.com/kryptogo/kyc-%E8%88%87-aml-%E7%9A%84%E5%B7%AE%E5%88%A5%E5%88%B0%E5%BA%95%E6%98%AF%E4%BB%80%E9%BA%BC-c65dac86e3ac>

17 資料來源：天下雜誌網頁版 <https://www.cw.com.tw/article/5077992>



另外，其「法遵管理儀表板」，以文字探勘技術，計算關鍵字於監管規範中出現的頻率，以「文字雲視覺化」的方式，觀察於指定範圍中較顯著的字詞，好瞭解法規異動趨勢，並以智能方式對主管機關裁罰案件進行分析，察知主管機關的監理重點，發掘更深一層的法遵風險變化，以便採取有效的管理措施。¹⁸

案例（九）：玉山銀行使用 AI 做風險管控

在 2021 年 Fintech Taipei 的玉山銀行的時任數位長唐枏演講中，玉山銀行揭露了自己在風控 AI 上有五大應用，分別是警示帳戶提醒、信用卡盜刷偵測、AML¹⁹ 監控、承作價值模型，以及智能反詐騙平台。底下針對玉山金控對媒體揭露的信用卡盜刷偵測、AML 監控做深入說明。

信用卡盜刷，一直是銀行界想要防範但是勞心勞力卻效果不彰。對玉山銀行而言，不只有數百萬張信用卡在市面流通，每天還得處理數十萬次的刷卡交易。如何從這些交易中揪出異常行為、即時攔截，是他們一直面對的難題。玉山銀行便因此發起一項人工智慧專案，要用刷卡歷史資料，訓練一套信用卡盜刷 AI 偵測模型，好能依此判斷盜刷風險，且提高辨識準確度。最後模型完成，計算風險。再將這筆交易的盜刷機率值，回傳給前端信用卡處系統，以決定交易是否終止，並依據風險程度，來啟動如電話通知、簡訊通知等相對應措施。這套 AI 信用卡盜刷偵測模型在 2019 年正式上線，現在模型可以在 0.1 秒內，就能回傳盜刷風險。玉山銀行內部的統計，這套模型在 2020 年替玉山阻擋了上億元的風險損失。²⁰

AML 偵測對銀行業非常重要，玉山銀行靠 AI 模型和三大步驟，來簡化很耗工時與人力的 AML 負面新聞的蒐集工作。其 AML 黑名單偵測流程如下，以自然語言處理（NLP）技術來加速負面新聞辨識，將流程分為負面新聞分類、事件聚合分析、黑名單資訊擷取三大步驟。最後呈現在資訊於前端介面，讓同仁可直接編輯、驗證，大大節省人力跟時間。²¹

18 資料來源：鉅亨網 <https://news.cnyes.com/news/id/4743495>

19 AML, Anti-Money Laundering，中文翻成反洗錢，是指金融機構和政府用於預防和打擊金融犯罪（尤其是洗錢和恐怖主義融資）的總體，更廣泛的措施和過程。資料來源：<https://medium.com/kryptogo/kyc-%E8%88%87-aml-%E7%9A%84%E5%B7%AE%E5%88%A5%E5%88%B0%E5%BA%95%E6%98%AF%E4%BB%80%E9%BA%BC-c65dac86e3ac>

20 資料來源：ITHome <https://www.ithome.com.tw/news/146199>

21 資料來源：ITHome <https://www.ithome.com.tw/news/146201>

在警示帳戶提醒上，玉山銀行是在資料庫端，利用 AI 針對帳戶的過往活動紀錄，對銀行提出警示，精準度就提升了 40 倍之多，有效降低營運風險。²²

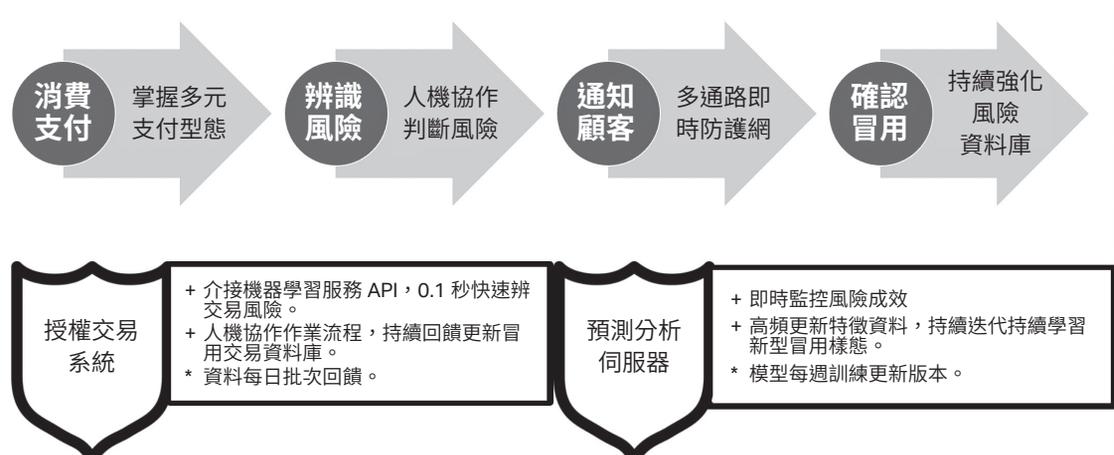


圖 6.4 玉山銀行的智能風控平台

圖片資訊來源：2021 數位新經濟線上論壇玉山銀行 Fintech 後疫情時代以數據驅動掌握行銷致勝關鍵簡報，圖裴有恆製作

案例（十）：富邦金控的人工智慧防詐系統

富邦金控在人工智慧防詐有多個系統，包含防止詐騙案的「智能防詐生態圈」、防止企業識別詐欺的「偽冒網站暨行動軟體 App 偵測及防禦機制」，以及防止金融產品詐騙的「智能風險評估系統」。

富邦金控與刑事警察局 2021 年底開始合作「AI 智能防詐模型」，與台北市警局合作「智能防詐生態圈」是其 AI 防詐的利器：在一般詐騙案件中，通常會有兩個重要的金流關鍵點，第一個是引導被害人把錢轉入人頭帳戶，第二個則是透過車手把人頭帳戶的錢提領出來，而「智能防詐生態圈」就是提前設定攔阻機制，增加阻詐成功率。

轉帳方面的攔阻機制是以大範圍的歷史詐騙態樣建立 AI 智能模型，陸續在台北富邦銀行（簡稱北富銀）分行櫃台、網路銀行及 ATM 提款機等服務節點增設提醒機制，協助民眾及早警覺可疑交易，避免受騙匯出款項。

²² 資料來源：MoneyDJ 理財網 <https://www.moneydj.com/kmdj/news/newviewer.aspx?a=84eee95a-e57a-49f0-a8fd-b3c27c2eef89>