

國立臺灣大學理學院心理學系

碩士論文

Department of Psychology

College of Science

National Taiwan University

Master's Thesis



線上學習成效的關鍵指標

Leading Indicators of Online Learning Performance

沈家齊

Chia-Chi Shen

指導教授：黃從仁 博士

Advisor: Tsung-Ren Huang, Ph.D.

中華民國 114 年 7 月

July 2025

## 誌謝



首先想謝謝我的指導教授黃從仁老師，我在大三加入從仁老師的實驗室，是我對心理資訊學領域的啟蒙導師。老師帶領幾乎沒有基礎的我參與不同的專案，開啟我對資訊領域的興趣，更支持我在碩士期間到外商實習進修。我一直記得老師常說的：

「actual competence 不等於 perceived competence」，在各種會議或簡報中，這句話帶給我很多幫助。從仁老師也在論文撰寫的過程中給我許多寶貴建議和鼓勵。此外，謝謝王雅鈴和蔡欣穆教授擔任我的口試委員，給予我相當多有建設性的建議，並讓我更加清楚自己論文的不足之處，在口試過程中獲益良多。可以說真正讀完碩士兩年，一步一步摸索建立自己研究的過程中，才終於能認為自己對做研究的道理有了基本認識。

在實驗室的這四年也很感謝有昱維、孟暉、昌育和中婷的指點和幫助，還有很強的夥伴彥廷和畢業救星柏圓、一起收實驗的翔易，以及統計學程的大家。這段時間在生活上，很感謝我的家人給我鼓勵和關懷，謝謝我溫柔的姐姐們，還有聽我講非常多心事、無條件支持我的靜玟、品樺、子瑜、芊樺、孝媛、鳳君，很照顧我、給我情緒支持的明彤，常一起吃飯大聊的子琿，陪我過生日的意婕，一起快樂出去玩的地調讀書會，溫暖的人資小夥伴，每次見到都還是一樣可愛的書班的大家，貓奴郁安。還有和心理系的大家，無論是一起跳舞或吃大餐，都是我很珍惜的快樂放風時光。還有其他許多帶給我能量的朋友們，謝謝你們在這段過程中給我的包容和鼓勵。

在工作上，非常幸運能在微軟實習，認識 A 組這一群友善幽默的人，謝謝 Angela 當最有紀律的人，論文幫 Polo, Eden 和 Lewis 這麼多個晚上跟我一起留守辦公室趕論文，James 傳授我很多技術知識，Anderlin 揪大家出國，Claire 是我的吉伊卡哇夥伴！也很感謝在 Azure Maps 的正職同事們都很好相處，在這裡學到很多、進步很多，是這一年很珍貴的收穫。

最後感謝台大，讓我在這裡度過充滿可能性的六年，也謝謝自己好好體驗這段時光。



## 摘要

本研究旨在深入探討影響線上學習成效的關鍵指標，研究目的在於綜合性探討關於過去研究中常見學習活動、行為及著名的間隔效應（Spacing Effect）與學生學業表現的關聯。本研究進一步將學習行為分為「資源瀏覽行為」（Navigation）、「知識獲取行為」（Input）與「知識應用行為」（Output），以求更細緻地探討這些行為的次數、行為分布及學習間隔對學業成效的影響。

首先，研究發現學生的早期表現，特別是在課程初期和中期的作業得分，對最終成績具有高度預測力，其影響力與課程後期的得分相當，強調了在遠距學習環境中進行早期干預的必要性。

另外，結果顯示學習行為的間隔對學業成效具有顯著的預測能力。其中，作業完成率及知識獲取行為的點擊頻率為負向因子，而知識獲取行為次數的波動程度以及知識運用行為的間隔天數則呈現正向影響，顯示高成效學生可能展現出更具策略性的學習行為，諸如適當的延長學習行為間隔、選擇性的作業繳交和資源投入。此外，資源瀏覽行為的平均間隔天數過長則顯示出負向預測作用，推測是因瀏覽行為實屬策略規劃行為的範疇，而非實質的學習活動，因此同樣體現了規劃學習策略之行為對最終成績表現的重要性。

總體而言，本研究呼籲針對線上學習的早期預警方法以及學生策略規劃行為的關注，強調未來研究應側重遠距教學場域的學習策略，並建議在教學實踐中加強對學習行為的監測與分析。

**關鍵字：**線上學習、學習成效、學習行為、早期干預、學習分析、混合效應模型、類神經網路、間隔效應、學習策略

# Leading Indicators of Online Learning Performance

Chia-Chi Shen



## Abstract

This study aims to explore the key indicators influencing online learning outcomes by comprehensively analyzing the relationship between common learning activities, behaviors, and the well-known Spacing Effect, as discussed in previous research. To further investigate these relationships, the study categorizes learning behaviors into "Navigation," "Input," and "Output" behaviors, examining the frequency, distribution, and spacing intervals of these behaviors and their impact on academic performance.

The findings reveal that students' early performance, particularly in the initial and middle stages of the course, strongly predicts their final grades, with its predictive power comparable to that of late-stage performance. This underscores the importance of early intervention in online learning environments. Specifically, early performance serves as an effective early warning indicator, enabling the identification of at-risk students for timely support.

Moreover, the study shows that the spacing of learning behaviors significantly predicts academic performance. Lower task completion rates may correlate with better performance, while longer intervals between Output behaviors and greater fluctuations in Input behaviors positively affect learning results. High-performing students may exhibit more strategic learning behaviors, such as maintaining appropriate learning intervals, selectively submitting tasks, and investing focused effort. On the other hand, excessively long intervals between Navigation behaviors negatively predict performance, suggesting that such behaviors, more related to learning strategy planning than actual learning, are important for academic success.

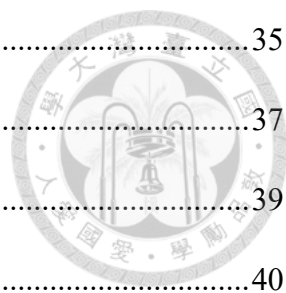
Overall, this study emphasizes the need for early warning systems and attention to learning strategies in online learning environments, suggesting that future research should focus on exploring learning strategies in distance education and improving the monitoring and analysis of learning behaviors.

*Keywords:* Online learning, learning outcomes, learning behavior, Spacing Effect, early intervention, learning strategies

# 目次



誌謝.....	II
中文摘要.....	III
英文摘要.....	IV
目次.....	VI
表次.....	VII
圖次.....	VII
第一章 緒論.....	1
第一節 研究背景.....	1
第二節 研究動機.....	6
第三節 研究目的.....	9
第二章 研究方法.....	10
第一節 資料集.....	10
第二節 資料前處理與樣本篩選.....	15
第三節 建模分析.....	17
第三章 研究結果.....	22
第一節 資料分布.....	22
第二節 混合效應模型.....	24
第三節 類神經網路模型.....	29
第四章 綜合討論.....	35



第一節 研究發現與貢獻.....	35
第二節 研究限制.....	37
第三節 未來展望.....	39
第四節 研究結論.....	40
參考文獻.....	42
附錄.....	48

## 表次

表 1：混合效應模型分析結果.....	25
---------------------	----

## 圖次

圖 1：資料集內人口變項分布圖.....	23
圖 2：Base Model 之 SHAP 值分析結果圖 .....	31
圖 3：Full Model 之 SHAP 值分析結果圖 .....	33



# 第一章 緒論

## 第一節 研究背景

### 壹、線上學習的興起

#### 一、高等教育與終身學習

在全球化與知識經濟高度發展的當代社會中，學習已不再侷限於特定年齡層或單一教育階段，而是成為個體持續適應變動環境、累積人力資本與提升社會參與能力的核心能力。根據聯合國教科文組織（UNESCO）提出的終身學習理念，教育應視為一種橫跨整个人生歷程的過程，涵蓋正式、非正式與非制度化的各種學習形式，旨在促進個人整體發展與社會可持續性（Tawil & Locatelli, 2015）。

在此背景下，高等教育作為終身學習體系中的關鍵節點，肩負促進知識創新、提升社會整體素養與培育具全球競爭力人才的功能。隨著科技變革速度加快與就業市場需求轉變，傳統的教育模式正逐漸轉向更具彈性、開放與學習者中心的取向。許多成人學習者選擇重返校園或參與線上課程，以因應職涯轉換、技能更新或個人發展需求，顯示高等教育正在朝向更多元、去中心化的樣貌發展（Bélanger & Tuijnman, 1997）。

隨著知識經濟與資訊科技的快速發展，高等教育的角色與型態正經歷深刻轉變。傳統以教室為核心的教學模式逐漸鬆動，教與學的界線從特定時間與場域延伸至數位空間。脫離學校體系後，成人學習者對彈性、自主與個別化的需求日益提升。許多高等教育機構紛紛導入數位教學平台與混成學習模式，以回應不斷變動的學習型態與教育挑戰。

## 二、疫情催化遠距教學的普及

這場變革在 2020 年 COVID-19 疫情爆發後被劇烈加速。各級學校與大學因應防疫需求大量轉向遠距教學，使線上學習從選項成為主流。教育工作者與學習者必須迅速適應虛擬學習環境（Virtual Learning Environment, VLE）的操作與互動模式，並在缺乏實體教室互動的情境中重新建構教學流程與評量標準。學生不再受到課堂物理空間的規範，取而代之的是必須自主掌握進度與內容的開放學習環境。教師則失去了傳統教室中觀察學生反應與即時調整教學的機會，對學生的理解與掌握程度只能依賴平台紀錄與作業表現間接推估。

這樣的變化帶來前所未有的挑戰，也創造了新的契機。隨著學習歷程全面數位化，學生的行為在學習平台上留下大量可觀測的操作資料，例如資源點擊、登入頻率、作業提交時機等，這些數位痕跡（digital trace data）逐漸被視為理解學習歷程與預測學習成果的重要資訊來源。特別是在缺乏傳統互動與監督的遠距教學場域中，教育者更迫切需要依賴這些行為資料，以掌握學生狀態、辨識學習風險，並設計適時而個別化的教學干預。

## 三、以線上平台做為學習研究媒介

隨著學習歷程的全面數位化，學生在虛擬學習環境中的各種行為都被即時記錄並轉化為可分析的數據，形成所謂的數位痕跡（digital traces）。這些資料不僅涵蓋操作層次（如點擊次數、登入頻率、作業提交時間），也可能反映了學生在學習過程中潛在的參與節奏、策略使用與資源調節能力。相比傳統教育場域，在缺乏面對面觀察的線上學習情境中，這些客觀而細緻的行為紀錄，成為理解學習狀態與成效的關鍵資訊來源。當代教育政策與學術研究亦日益強調「學習能力」（learning to learn）作為現代公民應具備的基本素養，並將學習視為應對資訊爆炸、人工智慧與全球不確定性的關鍵策略。由此出發，如何捕捉學習歷程、評估其效率，並預測學習結果，成為當代教育心理學研究的核心議題。

在線上學習環境蓬勃發展的背景下，學習分析（Learning Analytics）與教育資料探勘（Educational Data Mining）等方法應運而生，並逐漸成為連結教育實務與數據科學的交叉領域。這些方法能夠系統性地擷取並分析大量學習行為資料，進行預測、分群、風險辨識與成效追蹤，協助教育者洞察學習者的狀態與潛在困難，進而提供更具針對性的教學設計與即時干預。

線上平台之所以特別適合作為學習研究的媒介，不僅在於其資料豐富性與可取得性，更在於其結構化與高頻率的記錄特性，使研究者得以跨時間、跨課堂以多維度分析學習行為。此外，平台紀錄具有高度客觀性，避免了傳統測驗或自評工具中可能出現的反應偏誤與回溯誤差。透過這些資料，研究者不再僅能仰賴終點性的成績或靜態問卷結果，而能真正深入探索學習的過程本身，觀察學習是「如何發生的」、學習者「如何做出反應」、以及「哪些行為模式具有預測意義」。

總結而言，線上學習平台不僅是數位時代教學實踐的重要載體，更已成為蒐集、分析與理解學習歷程的關鍵研究場域。特別是在疫情加速線上教學普及的當下，這些資源不僅支撐教學活動的持續運行，也為學術界提供了深入探索學習行為與成效關聯的全新可能。線上教學的興起不只是形式的改變，更為教育研究開啟了一條理解學習本質與優化教學介入的新徑路。

## 貳、學習行為與成效之關聯

隨著數位學習平台的廣泛應用與資料可得性提升，學習行為成為教育研究中備受關注的分析對象。不同於傳統研究主要依賴問卷調查或教師觀察作為學習歷程的描述依據，研究者得以透過學習管理系統（Learning Management Systems, LMS）即時且客觀地紀錄學生的學習行為，嘗試理解其與學習成效之間的潛在關聯，以補足自我報告資料的主觀偏誤。

本節將從現有文獻出發，探討不同類型的線上學習行為與學習成效之間的關聯，並說明其背後可能涉及的學習機制與理論基礎，作為後續研究設計與變項選擇的理論依據。

### 一、學習行為與課堂參與度

學生於課堂參與之活躍程度與學習表現密切相關，許多研究中學者試圖以不同種類的行為資料捕捉學生的學習投入程度。學生對課堂的參與度高、展現積極的學習行為，往往與其學業成績呈正向關係（Credé 等, 2010; Li & Xue, 2023）。

在遠距學習的情境中，學生的出席率、作業遲交與缺交情形亦常被用以觀察其學習參與度，進而評估學習表現與其自我調控能力之間的關聯。Morris 等人（2005）的研究中則指出缺交作業數量是能顯著預測學業成績的行為指標之一，其與總成績呈現

顯著負相關。Y. Wang 等人 (2025) 所提出的行為參與評量架構中，也將線上作業提交率納為五大面向中「專注」一項的指標之一，並指出其與學習成效具有顯著相關性，顯示準時完成任務的行為有助於學業成就的累積與穩定。You (2016) 提出的研究設計以線上課程的準時出席率作為自我調節學習的行為指標，結果顯示規律的學習行為（意即準時出席率）是最顯著的成效預測因子，其次為準時繳交作業與分段學習次數。進一步地，Kizilcec 等人 (2017) 在對六門大型開放式線上課程 (MOOCs) 進行行為分析時，指出時間管理為有效預測課程完成與成就的重要自我調控學習變項。Kizilcec 和 Halawa 於 2015 的研究亦指出，儘管線上課程為學生自主安排時間進行且不設固定截止日，但時間管理不佳仍是最常見的中輟 (disengagement) 原因。多數學習者雖表示有完成所有作業的意圖，然而實際上僅有少數能真正達成此目標，顯示缺交與遲交現象在遠距課程中是明顯存在的議題，並能有效地反映學生的課堂參與程度。

得益於線上學習管理系統提供的資料豐富度，學者亦積極利用其他精細的行為數據及豐富的操作行為來衡量學生的參與程度，例如登入頻率、資源瀏覽次數、作業完成率、討論區互動行為等，來探討學習參與度與學習成效之間的關聯 (Joksimović 等, 2018)。Wang 和 Newlin (2002) 的研究指出學生的自我效能感不僅與其選擇遠距課程的動機有關，也與其線上互動頻率及最終學習表現具顯著關聯。Y. Wang 等人 (2025) 則建立了一套包含五大面向、共十六項指標的行為參與評量架構，並發現學生長時間進行線上學習與頻繁瀏覽學習資源等行為，能顯著預測其學業表現。

然而部分研究指出，學生在線上進行的學習操作頻率或次數雖能部分捕捉學生的參與程度，但僅對學習表現的預測有部分貢獻，顯示學習行為頻率對成效的影響或許隱含更複雜的關聯性。Morris 等人 (2005) 分析學生在多門大學線上課程中的操作紀錄，指出成功完成課程的學生，其線上閱讀教材與瀏覽討論的次數與時間均顯著高於未完成者，然這些參與指標僅能解釋學生成績的三成。Johnson (2005) 以混合式課程為例，從點閱次數、閱讀文章數與張貼文章數三個指標量化學生的線上行為，得出的研究結果顯示適度的發文與點閱次數可提升事實性與應用性題目的表現，但過高的互動反而可能分散學習焦點，造成成效下降。

同樣值得注意的是，線上學習時長對學習表現的影響力也與一般預期呈現落差，並非呈現單純線性的正向影響。Macfadyen 及 Dawson (2010) 以加拿大一所大學的全線上課程為對象，蒐集並分析學生在學習平台上的多項行為指標，嘗試建立一個能

預測學業表現的早期預警模型。他們分析的變項涵蓋點擊頻率、討論參與、評量完成情況等。最終研究結果顯示，總討論貼文數為最具預測力的變項，而學習時數與學業成績的關聯性卻極低，並未成為有效的預測因子。類似的結果在 You (2016) 的研究中也可以觀察到，其分析結果顯示單純的總觀看時間並不具預測力。因此，欲評估學生「線上學習參與度」時，應審慎設立合適的操作型定義。

總體而言，線上學習環境所提供的高度可追蹤性，使學習者與資源之間的互動不再是一個模糊的抽象歷程，而是一組可觀察、可量化且可解釋的行為數據。這不僅擴展了教學評量的範疇，也為學習成效預測與即時干預策略的設計提供了新的依據。在此脈絡下，出席率及互動頻率等變項雖仍為學習投入的重要指標，但其意義需與其他行為變項一同解析，方能掌握學習歷程的全貌。

## 二、學習行為的時間分布與節奏

學習行為的頻率並非影響學生表現的唯一關鍵，仍仰賴其他因素共同解釋，因而「學習時間的分布」也是探討學習策略與行為研究時的熱門議題之一。學習行為的分布形態一般被分為兩種，集中式學習 (massed learning) 指的是在一段時間內持續、密集地針對相同主題進行學習；分散式學習 (distributed learning) 則指涉相反的概念，強調將學習時程拆分成多個部分並加入間隔時間，間隔期間可以進行其他主題的學習或活動。

分散式學習所帶來的好處實際上應歸功於間隔效應 (spacing effect)。間隔效應指的是當單一刺激被重複學習時，把數次學習分散於適度間隔的時間點，比起連續密集學習，更能強化並持久保存記憶，尤其使用延遲測驗 (delay test) 的效果更顯著 (Jacoby, 1978)。Cepeda 等人 (2006) 彙整 184 篇文章、317 個實驗，計 14 000 多名受試者的成績，證實在控制總學習時間不變的條件下，只要將重複學習拆成兩次以上，最終回憶表現皆優於完全連續呈現。該研究同時探討學習間隔 (Inter-Study Interval, ISI) 及保持間隔 (Retention Interval, RI，即最後一個學習階段與測驗間的時間間隔)，發現間隔效應的優勢從不足 1 分鐘的保持間隔到超過 30 天皆未消失。這意味分散優勢並非僅在長期測驗才現身，而是橫跨短、長保持間隔的普遍現象。值得注意的是，間隔並非越長越好。在固定的保持間隔下，學習間隔與回憶 (recall) 表現呈現倒 U 型關係，ISI 從極短逐步拉長時，學生的表現會持續上升，然當 ISI 超過某臨界點又開始下滑。

研究者也嘗試將間隔效應應用於不同的教育領域中，舉凡外語、科學及數學領域都有相關研究證實了間隔效應的效果 (Bahrick 等, 1993; Kim & Webb, 2022; McDaniel 等, 2013; Reynolds & Glaser, 1964; Schutte 等, 2015; Vlach 等, 2022)。

Carvalho 等人 (2020) 則進一步探討在線上教學的環境中，學習者是否會自發性地採用分散式學習策略，以及此類自我調節行為是否與實際的學習表現有關。他們利用一門大型開放式線上心理學課程中的學習紀錄資料，分析學生在整個學期中完成每個單元所花費的學習會話次數，作為衡量間隔行為的指標。研究結果顯示，學習者若能將學習時間分散在多次會話中，即便總花費時間不變，其單元測驗與期末考的表現通常較佳。換言之，即使間隔行為並非由研究者控制，而是學習者自行決定，仍可觀察到明顯的間隔效應。

此外，該研究進一步指出，不同能力與行為特徵的學生在間隔行為上的差異。高能力學生傾向於更頻繁地分散學習時間，但間隔策略的效益卻在低能力學生身上更為顯著，可能因為間隔能夠部分補償其缺乏主動練習等其他有效策略的劣勢。類似地，當學生未積極參與單元活動時，若能將學習過程拉長至多次會話，其學習成效仍有提升，顯示間隔行為可能是一種具有彈性的替代學習策略。

Carvalho 等人的研究提供了重要的實證基礎，證明間隔效應不僅存在於實驗室控制條件下，也能在真實且具高度自我調節特性的線上學習情境中發揮作用，並強調了學習行為差異在策略使用與效果之間所扮演的關鍵角色。

## 第二節 研究動機

### 壹、基礎行為指標不足以預測學習成效

在過去的研究中，許多學者利用不同的學習行為指標來探討學生參與度與學習成效之間的關聯，然而，這些變項的解釋性並未能充分捕捉學習過程中的複雜性。例如前一節提到的「學習時長」這一變項，在實證研究中被指出與學業成績之間的關聯性非常微弱 (Macfadyen & Dawson, 2010; You, 2016)。另亦有研究明確指出「學習事件的發生頻率」和學生真實的課程參與度實無明顯關聯 (Hadwin 等, 2007; Misanchuk &

Schwier, 1992)。

再者，間隔效應雖為傳統教育研究中的重要議題，對學習成效有顯著的正向影響，卻較少在線上教學的場域中被討論。過去的研究通常將焦點單獨放在行為頻率或作業繳交情況上，未能綜合考慮這兩種學習行為的特徵，研究設計中也經常忽略了學習行為的間隔性這一重要指標。

由此可見，若能結合多種行為變項，同時探討學習行為的頻率與時間分布，從而整合性地完整捕捉學生們複雜的學習型態，以及進一步探討這些變項如何共同影響學習成效，將能有效填補現有線上教育研究的重要缺口。

## 貳、間隔效應的研究缺口

### 一、缺乏學習行為的分類探討

在先前的討論中，提及許多學者利用學生的點擊行為、瀏覽次數或時間分布來預測學習成效，然而這些研究中皆未對學習行為進行更細緻的分類分析，主要是由於資料量的限制，難以深入細分每種行為對學習成效的影響。根據 Moore (1989) 與 Hillman 等人 (1994) 的互動理論，學習行為可視為學習者與內容、教師、同儕或系統介面之間的互動模式。而近期的研究，如 Qiu 等人 (2022) 提出的行為分類預測框架，以及 Ye 等人 (2022) 針對學生行為特徵融合與選擇的方法，進一步根據線上學習流程將行為區分為「準備」、「吸收」、「互動」、「鞏固」等階段，詮釋學習歷程的階段性。這些研究將學習行為細分為不同的類別，並成功利用分類提高學習成效預測的準確性。

學習行為的分類優勢在於，其不僅能讓我們更精確地識別學生的學習狀態，還能幫助我們發現學習過程中不同行為間的複雜關聯。這種分類方法有助於揭示學生在不同學習階段的需求和合適的策略，並為教師提供具體的反饋和指導。

因此，本研究引用類似的行為分類方法，並特別針對點擊行為時間分布的分析。由於過去關於間隔效應的研究鮮少以不同類型的學習資源切入，探討各項分類是否適用不同的時間分配方式。透過將學習行為分類，我們可以更細緻地理解學生在學習過程中的各種行為模式及其對學習成效的個別影響。例如，學習準備行為、知識獲取行為等各類型的行為可能在時間分配上有著顯著差異。這些差異在預測學習成效時可能

會產生重要的影響，因而有望在現有的學習成效預測模型基礎上，提供一個更加細緻且具有實用價值的分析框架，進一步提升預測的準確性以及對學習成效的解釋性。

## 二、應用於高等教育的實證研究不足

在既有間隔效應的相關研究中，受試者通常會被要求針對單一刺激反覆學習及練習，並且主題也較為簡單，例如背誦相同的單字，或是重覆練習簡單的加法運算，因此通常也以學生的記憶及熟練度作為學習效果的評估目標。然而，這樣的學習及評量方式一般常見於基礎教育的場域中。一旦進入高等教育階段，教學對象的成熟度更高，教學的內容往往也會涉及更加複雜的知識及概念，導致過去的研究可能無法適用於這樣的教育情境。亦有研究顯示分散式學習會導致學習過程中的注意力減低，並且提高學習當下遺忘的表現 (Vlach 等, 2022)。

有鑑於線上學習平台在當代高等教育中日益成熟並扮演重要角色，本研究旨在針對高等教育情境進行實證探討，結合比較先前提到的多項學習行為變項，補足間隔效應在此領域中的研究缺失。

## 參、未能運用學期間評量的時間性資訊

作業及小考等評量工具作為課堂中的常客，亦為影響學生學業成就的重點因素之一，一般認為在課程中安排作業能夠有效提升學生最終的學業表現 (Grodner & Rupp, 2011; Latif & Miles, 2020)。根據 Eren 及 Henderson (2006) 的研究，不同學業成就的學生群體中，作業的效應並不一致。具體而言，高成就與低成就學生對作業的反應較為顯著，而中等成就學生則可能未能從額外的作業中獲益。另外，兩人也發現並非所有科目的學習表現都能透過作業提升 (Eren & Henderson, 2011)。

然而，多數現有的研究主要集中於評量的存在與否或評量數量對學習成效的影響，較少關注評量分數所提供的時間性資訊。與最終成績不同，作業與小考往往在整個課程中多次進行並分佈於不同的學習階段，因此能更精確地反映學生在特定時期的學習成效。從提前干預 (intervention) 的角度來看，這類時間性資訊具有極大的價值。如果能夠利用這些資訊識別出影響學生最終表現的關鍵學習階段，教育者便能即時針對有需求的學生進行幫助，進而提升學習成效。

因此，本研究同時探索學生在課程的不同階段（前、中、後）中的作業表現如何

影響最終的學習成效。透過對作業效應的階段性分析，我們期望能夠揭示不同學習階段的學生對作業的不同反應，從而為教育者提供更明確的建議。



### 第三節 研究目的

基於前述研究背景與動機，本研究提出更加全面且細緻的研究方法，以加深理解高等教育於線上教學場域中各項學習指標對學習成效的影響。

本研究一併考量過去研究中普遍認同對學習表現有部分解釋利的重要指標，包含作業繳交遲缺、作業成績、點擊頻率以及學習行為的分散程度，結合學習行為的分類方法，並善用作業成績提供的階段性資訊，以期能整合性地理解線上學習行為的多維特徵後，提出一個更為精確的預測模型。該模型有助於我們深入了解多項學習指標對學習成就的影響力，以及指標間的相對重要程度，並為線上教育提供者提供實證資訊，以實踐對學生學習狀況的即時監測與早期干預。



## 第二章 研究方法

為了回應第一章提及的研究缺口，本研究採用線上高等教育之相關資料進行分析，並提出更新的研究架構，以深入理解線上學習行為與學業成就的複雜關聯。

本章共分為三節，首先介紹本研究採用的資料集，接著為了符合本研究的研究目的，對資料集作相應的篩選及合適的前處理，最後一節則講述本研究運用的兩種分析方式。

### 第一節 資料集

#### 壹、資料集簡介

本研究所使用之資料集為 Open University Learning Analytics Dataset (簡稱 OULAD) (Kuzilek 等, 2017)，由英國開放大學 (Open University, OU) 公開釋出，旨在支援學習分析 (Learning Analytics) 領域之研究。其建置過程可分為三個主要階段：資料收集、資料篩選與資料匿名化處理。資料最初由英國開放大學以 SAS 技術建置之資料倉儲中萃取，內容包含自 2013 年至 2014 年間學生在虛擬學習環境 (Virtual Learning Environment, VLE) 中的學習行為紀錄，並結合學生之人口統計資料與學業表現資訊。資料中不同的課程稱為模組 (module) 依照開課年份與月份予以標示，每一模組可於不同時間重複開設，每次開設稱為一個學期 (module-presentation)。資料去識別化過程依據倫理與隱私原則進行，並運用 ARX 匿名工

具，確保個體無法被辨識。該資料集已通過 Open Data Institute 認證，並以 CC-BY

#### 4.0 授權方式開放使用。

OULAD 的最大特點在於其整合了多樣化的資料類型，包括學生的基本背景（如性別、年齡區間、教育程度、地理區域、剝奪指數等）、課程註冊紀錄、各項評量成績，以及學生每日於 VLE 上的點擊活動數據。其中，點擊活動依據點擊目標不同細分為 20 種類別，諸如課程首頁（homepage）、小考（quiz）和討論區（forumng）等。資料涵蓋共計 22 門課程、32,593 位學生，以及超過一千萬筆互動紀錄，為目前學界中公開可得且結構完整的學習分析資料集之一。

以下僅羅列及簡介本研究使用的變項，詳細篩選流程請參考第二節：

（一）地區（region）：標示學生居住的區域，諸如 London Region、East Midlands Region 等。

（二）教育背景（highest\_education）：修課學生的教育程度，以英國學制區分。

（三）IMD 指標（imd\_band）：英國多重剝奪指標，亦稱貧窮指標，反映學生居住區域的經濟狀況。

（四）性別（gender）：將學生區分為男性（M）、女性（F）。

（五）年齡區間（age\_band）：學生之年齡區間，諸如 0-35 歲、35-55 歲等。

（六）累積修習學分（studied\_credits）：學生當學期修習的總學分數。

（七）註冊日期（date\_registration）：學生註冊課程的日期，以整數標示相對於開課日經過的天數，0 為開課當天，正數為開課後，負數為開課前。



## 貳、學習成效指標

在學習分析相關研究中，選擇合適的學習成效指標是相當關鍵的步驟。OULAD 資料集中有一項較頻繁地被相關研究 (Qiu 等, 2022; Ye 等, 2022) 使用的學習成效預測目標變項 `final_result`，為學生在課程結束後的最終表現評估，分為優異 (distinction)、通過 (pass)、不通過 (fail)、退出 (withdraw) 四個種類。此變項為 OULAD 資料集中根據學生於整個學期中綜合表現所計算出的總結性指標。雖然 `final_result` 在多數研究中被視為有效的學習成效指標，但其本身係以學生整學期之評量結果為基礎計算，若使用學期間的作業成績或其他評量相關指標來預測此最終結果，可能導致預測模型在訓練階段即已隱含目標答案，而使我們對於各項指標的重要性做出帶有偏差的詮釋。為避免此一問題，本研究改以 `assessment` 資料表中類型 (`assessment_type`) 為 Exam 之成績 (score) 作為學習成效的主要衡量指標。該成績代表課程結束後之期末考試表現，為獨立計算、不涉及其他評量的成績，也是一項整合性的評估指標，能作為學生整體學習成果的有效指標。同時，為確保選擇的課程內容適合以考試形式作為最終表現的評估指標，本研究僅選擇 OULAD 中屬於 STEM (Science、Technology、Engineering、Mathematics) 領域且具期末考試成績之課程相關資料作為分析樣本。

此外，為降低連續變項在分類任務中可能產生的資料不平衡 (data imbalance) 問題，並提升模型對於不同表現學生之行為差異辨識能力，本研究將 Exam 成績自連續數值轉換為類別型變項，由分數低到高將資料筆數平均分為三組：「低分組」、「中分組」與「高分組」。然而在實務分析過程中發現，中分組樣本行為模式分布較為雜亂且無明顯變項與規律，因此本研究最終僅保留高分與低分兩組樣本進行二元分類預測任務，以強化模型對於極端表現學生間之學習行為差異辨識能力，以期模型之分析結果能更有助於我們深入掌握提升與降低學習成效的關鍵因素。

## 參、衍生變項設計

為提升對學生學習行為與表現之理解與預測能力，本研究在 OULAD 原始資料集的基礎上，進一步設計並建構多項衍生變項。這些變項依據學生的評量紀錄與虛擬學習環境 (Virtual Learning Environment, VLE) 中的點擊行為進行計算與彙總，旨在整合

性地捕捉學習歷程中時間性與行為面的差異，並提供模型更具解釋力的變項輸入。

## 一、課程識別變項

由於相同課程模組 (module) 可於不同學期 (presentation) 開設，為強化識別性與便於資料關聯整併，特將兩者組合為單一欄位 module\_presentation，作為課程呈現的唯一代碼 (例如 "FFF\_2013B")，用以對應學生所參與之具體課程時間單元。

## 二、與課程評量相關的衍生變項

為善用資料中的時間性資訊，衡量學生在不同時間階段的學習情況，本研究根據每一門課程的上課時長，將整體課程平均分為「早期」、「中期」與「後期」三個等長階段，並於每一階段計算平均分數。同時，本研究也參照過去研究，一併考量評量的完成率及遲交率。與評量相關的衍生變項包含以下：

(一) early\_score/mid\_score/late\_score：分別計算學生在課程早期、中期與後期所完成之評量的平均得分，反映學生在不同學習階段的表現趨勢。

(二) completion\_ratio：定義為學生實際提交之作業數與該課程所有應完成作業數之比例，為一介於 0 到 1 之間的數字，用以評估作業完成度。

(三) late\_ratio：為學生繳交評量時間晚於截止日的比例，為一介於 0 到 1 之間的數字，用以觀察其學習行為中之遲交傾向。

## 三、與點擊行為相關的衍生變項

OULAD 記錄的學生點擊行為係以日為單位結算，並且將點擊行為依據點擊目標類型細分為多項。誠如前述，由於本研究以學生期末考成績作為學習成效的主要評估指標，經過此條件的篩選流程後，適用資料中包含的點擊目標種類共計 13 種。在處理此類型學習行為資料時，由於許多點擊目標之間存在關聯性，或是在學生學習的過程中提供相近的用途。若直接使用平台所記錄的所有細項行為作為預測變項，不僅會導致變項維度過高、模型複雜度上升，並容易因為過度細緻的分類分散了潛在的高階變項對模型的貢獻，而降低整體結果的可解釋性。為此，許多相關研究提出將個別學習行為依據其在教學歷程中的功能加以分類與合併，以整合成更具代表性的行為變項群組。這樣的分類策略能幫助模型捕捉到資料中傳遞的整合性資訊，並使其預測結果更易於解讀，進而提供更高的應用價值。

兩篇近期以 OULAD 資料集為基礎的研究，分別提出了行為分類模型——Ye 等人 (2022) 建構的 E-learning Behavior Classification model (EBC model)，以及 Qiu 等人 (2022) 所提出的 Process-Behavior Classification model (PBC model)。這兩者皆根據 Sun (2015) 所整理的線上學習流程理論，將線上學習過程劃分為四個主要階段，並對應至以下四類行為類型：

(一) 學習準備行為 (Learning preparation behavior, LPB)：登入平台、開啟課程頁面等。

(二) 知識獲取行為 (Knowledge acquisition behavior, KAB)：瀏覽教材、觀看影片、下載資源等。

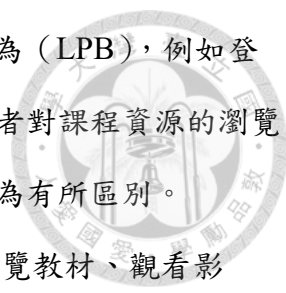
(三) 互動學習行為 (Interactive learning behavior, ILB)：參與論壇、討論、提問等。

(四) 學習鞏固行為 (Learning consolidation behavior, LCB)：填寫問卷、完成測驗、重複活動等。

此一分類方式並非僅以平台功能或互動對象為基礎 (Hillman 等, 1994; Moore, 1989)，而是強調學習歷程中不同階段對應的認知活動，具備更強的教育心理基礎與理論連結。Qiu 等人 (2022) 進一步證實，以此方式進行變項融合與分類建模，可以有效提升預測準確率與模型穩定性；Ye 等人 (2022) 則整合差分進化演算法與此類行為分類策略，提出具備自適應融合能力的模型，進一步提升模型效能與行為解釋力。

本研究基於上述觀點，對行為進行了更進一步的邏輯整合。具體而言，考慮到資料集中的 ILB 與 LCB 兩類行為樣本較少，我們認為可以將這兩類行為整合為「知識產出」行為類型，將其視為學習者在學習過程中的高階行為，主要反映其在後期學習過程中進行的互動、應用與鞏固等活動。此外，KAB 行為可作為「知識獲取」行為，代表學習者在學習初期進行的準備與知識吸收過程。

基於此邏輯，我們將原有的行為分類進行了如下調整，將行為分為以下三類，並以此作為後續模型建構的基礎變項：

- 
- (一) 資源瀏覽行為 (Navigation)：此類行為包含學習準備行為 (LPB)，例如登入平台、開啟課程頁面等。由於這些行為主要反映學習者對課程資源的瀏覽和探索，應視為學習的準備階段，與實際的知識學習行為有所區別。
- (二) 知識獲取行為 (Input)：此類行為屬於 KAB 類別，如瀏覽教材、觀看影片、下載資源等，反映學習者的知識吸收過程。
- (三) 知識應用行為 (Output)：整合 ILB 與 LCB，涵蓋學習者在學習後期利用所學進行的互動、應用等高階學習歷程。

此三分類架構保留原四類架構的理論基礎，同時結合任務導向與分析目的，兼顧資料簡化與學習解釋，有助於釐清其在整體行為變項中的中樞角色與代表性。基於此分類架構，本研究將於每一類別中計算下列行為變項：

(一) mean\_clicks：各類型資源之平均點擊次數，反映學生對該類資源的使用程度。由於 OULAD 僅記錄有點擊活動的天數，因此在計算步驟中會先取學生最早點擊日及最後點擊日，並將中間缺少點擊數據的天數補值為 0，再加總計算平均點擊次數。舉例而言，若 1 至 5 日中第 2、4 日各有 10 次點擊，則總點擊總數為  $0 + 10 + 0 + 10 + 0 = 20$ ，總天數為 5，得平均值為 4。

(二) click\_cv：點擊次數的變異係數，用以描述學生在不同學習日之間點擊次數的相對變化幅度，代表學習活躍表現的波動程度。其中計算變異係數所需的點擊次數標準差及點擊次數平均值，皆採用與前項 mean\_clicks 相同之補零計算方式。

(三) mean\_click\_interval：以同一學生針對該類資源的連續點擊時間間隔（天數）計算平均值，作為其使用頻率的時間指標，間隔越短表示使用越頻繁。若兩相鄰日期皆有點擊，則間隔為 0。

## 第二節 資料前處理與樣本篩選

本研究之分析目的為預測學生在課程期末考試中的學習表現，並探討其潛在影響因素。因此，資料處理階段以 OULAD 原始資料為基礎，結合衍生變項，依照研究目

標與變項品質進行嚴謹的資料清理與篩選作業，處理流程如下。



## 壹、樣本選取與學習成效分組

首先，根據 assessment 資料表篩選所有擁有期末考試成績 (assessment\_type = Exam) 的學生紀錄，作為本研究之主要預測目標變項。經彙整後共取得 4,959 筆有效樣本。為建立具區辨力的學習成效分類模型，進一步將期末考試成績以百分位分布劃分為「低分組」、「中分組」與「高分組」三類。為集中分析作業於明確表現差異之樣本，本研究排除中分組樣本，僅保留高分與低分兩組，最終樣本數為 3,375 筆。

## 貳、變項篩選與移除依據

在建立預測模型前，需針對原始變項進行適當篩選與處理，以排除無效、冗餘或對模型貢獻有限的欄位，具體依據如下。

### 一、常數欄位移除

is\_banked 為作業紀錄狀態指標，於本次篩選後樣本中值完全一致（全為 0），無變異性，因此不具資訊量，予以刪除。此外，weight 變項雖表示評量比重，但由於本研究僅分析期末考 (Exam) 成績，該評量比重在所有樣本中皆為 100%，不具區辨性，因此亦予以排除。

### 二、缺失值過多者移除

date\_unregistration (退選日期) 之缺失比例超過半數，不宜進行有效填補，故予以剔除。

### 三、資訊冗餘欄位移除

date (評量截止日) 與 date\_submitted (作業繳交日) 雖具時間資訊，但已透過衍生變項 late\_ratio 反映遲交行為，因此避免變項重複與模型偏重，予以排除。

### 四、變異性過低欄位移除

num\_of\_prev\_attempts 表示學生過去修習該課次數，然其標準差僅為 0.1，變異性極低，難以提供有效區辨，故亦不納入建模。



## 參、缺失資料處理

在上述變項過濾後，進一步檢查剩餘資料之完整性。由於本研究擬採用統計模型與類神經網路進行擬合，前者需完整樣本作為輸入，因此對資料中仍存在缺失值之觀測進行刪除處理。最終共保留 2,819 筆資料，佔前述樣本之 83.2%，作為本研究後續建模與分析之基礎資料集。

## 第三節 建模分析

為深入探討影響學生學習成效之因素與其關聯性，本研究針對 OULAD 資料集進行預測建模與變項解釋分析，並聚焦於預測學生期末考成績所屬之分數群（高分／低分），進一步探索具代表性的學習領先指標（leading indicators）。考量學習行為可能同時存在線性與非線性關係，本研究採用兩種具互補特性的分析方法：混和效應模型（mixed-effects model）與類神經網路模型（artificial neural network model），旨在建構具高度可解釋性的預測架構，以期對學習領先指標具備更全面且深入的理解。

上述兩種建模方法皆配合「雙重變項設計策略」進行，分別訓練兩種模型 Base Model 與 Full Model。Base Model 僅納入學生的基本人口統計變項，包括：gender、region、highest\_education、age\_band 以及合併後之課程時間代碼 module\_presentation，此為反映學生入學時可觀察之背景資料，作為基本參照基準。Full Model 則在 Base Model 基礎上進一步加入衍生變項與學習行為變項，例如點擊行為、作業完成率、階段性表現等，以完整模擬學生在學習過程中的互動與投入情形。

透過 Base 與 Full 兩版本模型於兩種方法下之比較，不僅可評估學習歷程相關變項對預測能力的貢獻程度，亦可辨識學生先天背景與後天行為間之關聯性與交互作用，進一步理解何種因素最具代表性並可作為潛在的領先指標，協助早期識別學習風險或提供個別化教學建議。

## 壹、混合效應模型 (Mixed effect model)



混合效應模型 (Mixed-Effects Model, 又稱階層線性模型 hierarchical linear model 或 multilevel model) 是一種統計建模技術, 提供良好的模型可解釋性與參數穩定性, 能有效呈現個別變項與學習成效間之線性關係, 特別適用於處理具有層級結構 (hierarchical structure) 或重複測量 (repeated measures) 的資料情境。此類模型同時納入固定效應 (fixed effects) 與隨機效應 (random effects), 能夠在捕捉整體變項影響的同時, 合理考量群體間的異質性, 避免傳統線性模型可能出現的偏誤估計。

考量到 OULAD 包含在不同學期修習不同課程的學生資訊, 本研究選用混合效應模型作為主要分析方法, 將此類群組變項作為隨機效應, 能有助於控制群間差異, 使跨群間的重要共同變項脫穎而出。

### 一、資料前處理

為避免變項間量級差異影響模型估計與收斂效率, 建模前將所有連續型變項進行 min-max 標準化, 使其數值範圍落於 0 至 1 之間。此處理方式可提升參數估計之可比性, 亦有助於後續與類神經網路模型進行變項重要性之比較。

### 二、建模流程

在本研究中, 固定效應部分涵蓋學生的基本人口統計變項與衍生的學習行為變項, 目的在於建構可解釋學生表現之主要預測因子; 隨機效應則涵蓋課程模組 (module) 與開課時段 (presentation), 此為 OULAD 資料中明確存在的群組層級, 透過建模可調整各模組間的系統性差異並控制潛在偏誤。如此設計可兼顧群內一致性與群間變異, 提升模型推論的穩定性與泛化能力。

實作上, 本研究使用 R 語言中的 lme4 套件進行模型擬合, 並配合 glmer() 函數構建二元邏輯斯迴歸的混合效應模型 (Generalized Linear Mixed Model, GLMM), 以學生是否為高分組為依變項。

整體而言, 混合效應模型提供良好的參數解釋性與結構穩定性, 能有效揭示學習行為變項與學習成效間之線性關係。然而, 其於捕捉高階交互作用或複雜非線性變項模式方面仍具限制。為補強此點, 後續分析則輔以類神經網路模型進行資料擬合與潛在結構發掘, 並以 SHAP 工具提升模型透明度, 形成本研究雙重分析策略的重要架構。



## 貳、類神經網路模型 (Artificial Neural Network Model)

儘管混合效應模型具備良好的解釋性與參數穩定性，能揭示線性關係並調整群組效應，其在處理高維變項交互作用與非線性行為模式時仍有其結構性限制。為補足此一不足，本研究進一步導入類神經網路 (Artificial neural network) 模型，期望藉由其高度的非線性建模能力與自動變項組合優勢，發掘潛藏於學生行為資料中的複雜模式與潛在學習結構。

### 一、模型架構與設計原則

本研究所使用之深度模型為參考 H2O 平台中常見的前饋式神經網路架構 (feedforward neural network)，並以 Keras 套件建構三層全連接神經網路 (Fully Connected Neural Network)。模型架構如下：輸入層依據處理後變項維度動態定義；兩層隱藏層，各含 50 個神經元，激活函數均採用 ReLU (Rectified Linear Unit)；每層加入 L1 正則化 ( $\lambda=1e-5$ )，以避免過擬合並促進變項選擇；輸出層為單一神經元，採用 sigmoid 函數以進行二元分類 (高分組 vs 低分組)；模型損失函數為 binary cross-entropy，最佳化方法為 Adam，並計算準確率 (Accuracy) 作為評估指標。

### 二、資料前處理與建模流程

在模型訓練前，首先將資料依變項 score (類別：高分組與低分組) 與所有預測變項分離。針對自變項資料進行自動類型判別，區分為數值型與類別型欄位，並建構前處理流程如下：

- (一) 數值變項經 min-max 標準化處理，使值落在 0-1 的範圍內。
- (二) 類別變項以 One-Hot 編碼進行轉換。
- (三) 所有前處理步驟與模型一併封裝進 sklearn 的 Pipeline 中，確保交叉驗證與實際預測時資料轉換一致。

### 三、模型評估

為確保模型之穩定性與泛化能力，本研究採用 5 折分層交叉驗證 (5-Fold Stratified Cross Validation)。每一折訓練皆重新執行資料前處理與模型擬合程序，並計算準確率 (Accuracy)，進行整體評估。最後彙整五折準確率的平均值與標準差，

以評估模型在不同資料切分下的一致性與表現。

#### 四、模型解釋與變項重要性分析

類神經網路模型雖然具備強大的非線性擬合能力，但其內部結構通常被視為「黑箱」，難以直接從參數層級解釋預測結果與變項之間的關係。為提升模型應用於教育領域的可信度與實用性，本研究導入解釋性人工智慧（Explainable Artificial Intelligence, XAI）方法，對模型預測進行細緻的後設分析，以了解各項變項對模型預測結果的貢獻方向與程度。

目前常見的 XAI 方法包括：

（一）Feature permutation importance：透過隨機打亂變項來觀察模型準確度變化，評估變項重要性，但無法解釋正負方向。

（二）Partial Dependence Plot (PDP)：繪製變項與預測值間的平均影響，但僅適用於單變項或雙變項，解釋力有限。

（三）LIME（Local Interpretable Model-Agnostic Explanations）：透過區域性擾動樣本並擬合簡單模型進行解釋，能提供局部預測依據，但無法判讀全域的變項正負貢獻。

（四）SHAP（SHapley Additive exPlanations）：基於賽局理論（game theory）中 Shapley value 概念，為目前較為完整且理論嚴謹的模型解釋方法。SHAP 方法針對每筆資料的預測結果，計算每個變項對最終輸出貢獻值的邊際增益。透過遍歷所有變項加入模型的順序組合，求出該變項在所有可能組合情境下的平均貢獻，最終形成具有方向與大小的 SHAP 值。其輸出形式為一組數值，反映某一變項在特定樣本中對預測結果造成的影響變化（提升或降低預測機率）。

本研究選擇採用 SHAP 作為深度模型的解釋工具，是以 SHAP 值能提供方向性的特點作為主要考量。SHAP 方法能指出每一變項對預測機率造成提升或下降的程度，方便量化每個變項的正負影響，這對於教育場域的應用特別重要，可辨識哪些行為變項對學習成果具有正向影響、哪些可能具有負面效應，因此特別適合本研究探討「哪些變項鼓勵學生表現提升」、「哪些行為與學習困難相關」等教育問題。此外，SHAP 不僅適用於單一樣本的個別預測解釋，也能整合出全體樣本的平均變項貢獻，有助於體現資料集的整體趨勢。

在實作上，本研究應用 DeepExplainer 套件針對訓練完成之 Keras 模型進行

SHAP 值計算，並配合轉換後的變項名稱與數據繪製圖表，直觀呈現變項重要性排序與其影響方向，以增進對模型決策邏輯之理解。





## 第三章 研究結果

本章共分為三節，第一節為資料分布，第二節講述使用混合效應模型進行的分析結果，同時比較 Base Model 與 Full Model 的結果異同。第三節為類神經網路模型的分析結果，並將 Base Model 及 Full Model 分為兩段，分別討論兩種模型的訓練結果及使用 SHAP 方法獲得的模型解釋。

### 第一節 資料分布

為更進一步理解建模資料集的組成特性，本研究針對訓練資料中主要人口背景與學習歷程變項進行分布分析。以下簡要說明資料集中人口背景變項之分布情形：

(一) 課程模組與學期 (module\_presentation)：樣本來自多個課程模組與學期開課呈現，其中以 *CCC\_2014B* 為最多。

(二) 地區 (region)：學生來自英國各地，分布相對均衡，前幾高比例為 *South Region*、*East Midlands Region* 等，呈現地域廣泛性。

(三) 教育背景 (highest\_education)：以具備 *A Level or Equivalent* 學歷者為最多 (超過 1,400 位)，*A Level* 指的是英國高中最後兩年 (Year 12 和 13) 的課程，是當地學生升讀大學所需的重要資格證明。其次為 *HE Qualification*，相當於大學以上的學歷。

(四) IMD 指標 (imd\_band)：即英國多重剝奪指標，分布均衡，未見極端偏態，顯示樣本來自不同社經地位背景。

(五) 性別 (gender)：樣本中男性 (M) 略多於女性 (F)，約占整體的六成，屬於性別不對稱分布，解讀分析結果時應注意此不均衡分布之潛在影響。

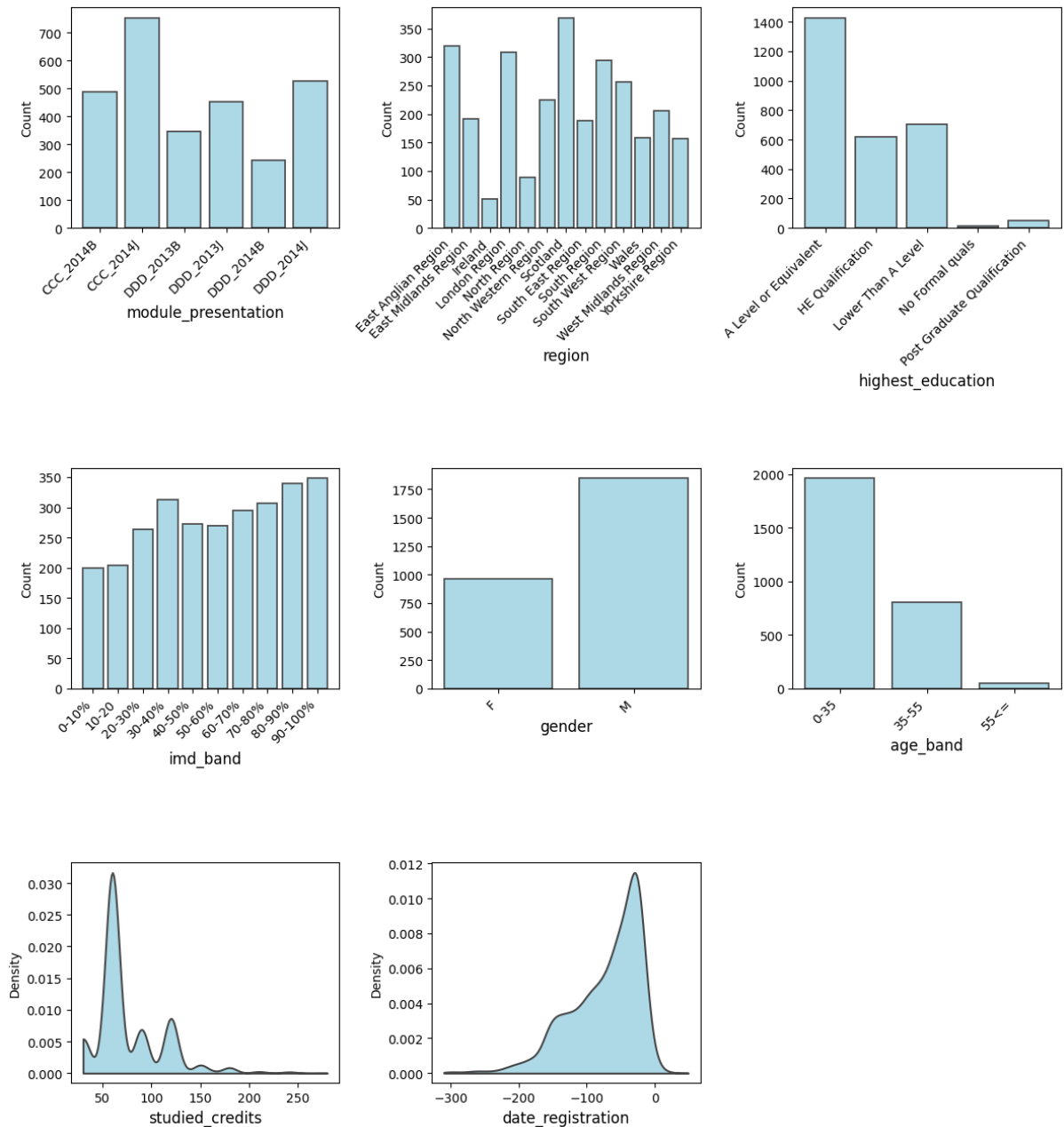
(六) 年齡區間 (age\_band)：以 0–35 歲學生為主，解讀分析結果時應注意此不均衡分布之潛在影響。

(七) 修習學分 (studied\_credits)：呈現右偏分布，多數學生修習學分集中於低至中間區間 (50~100)。



圖 1

資料集內各項人口變項分布圖



## 第二節 混合效應模型



本研究首先建構僅包含基本人口變項與課程代碼的 Base Model，作為預測學生高低學習成效的基準模型；進一步加入學習行為與表現等衍生變項後形成 Full Model，以比較不同變項層級對模型預測能力與重要變項的影響。

從模型解釋力觀察，Base Model 的邊際判定係數 (Marginal  $R^2$ ) 與條件判定係數 (Conditional  $R^2$ ) 分別為 0.091 與 0.173，顯示單以基本背景變項進行預測時，模型能解釋的變異有限。相較之下，Full Model 的 Marginal  $R^2$  與 Conditional  $R^2$  顯著提升至 0.713 / 0.796，說明當加入行為與表現變項後，模型整體解釋力大幅增強，特別是在固定效應部分的提升最為明顯。

在 Base Model 中，少數變項達顯著水準。居住地區為 *West Midlands* 者，其成功機率顯著較低 ( $\log\text{-odds} = -0.76, p < .001$ )；教育程度為 *Lower Than A Level* 相較於基準組 (*A Level*)，其表現亦顯著較低 ( $\log\text{-odds} = -0.71, p < .001$ )，而擁有 *Post Graduate Qualification* 者表現較高 ( $\log\text{-odds} = 1.38, p = .001$ )；年齡大於等於 55 歲之學生亦顯著偏向高分組 ( $\log\text{-odds} = 0.98, p = .010$ )；男性學生 (M) 相較女性更傾向落於高分組 ( $\log\text{-odds} = 0.17, p = .047$ )，惟效果不大，且需注意於前一節分析結果中曾提及男學生數量較女學生更多。

在 Full Model 中，除了前述變項中仍保有顯著性的地區 (*West Midlands*) 與教育層級 (*Lower Than A Level*) 外，更多行為與表現變項顯現出預測力。

首先，與學生評量相關的衍伸變項中，*early\_score*、*mid\_score* 與 *late\_score* 均為強烈正向預測因子，其中以中期成績 *mid\_score* 的效果最為顯著 ( $\log\text{-odds} = 7.48, p < .001$ )、*early\_score* 為其次 ( $\log\text{-odds} = 5.23, p < .001$ )、*late\_score* 最後 ( $\log\text{-odds} = 4.54, p < .001$ )。*completion ratio* 則為顯著負向預測變項 ( $\log\text{-odds} = -0.81, p = .053$ )，其與一般預期相反，或許暗示部分高分學生在作業提交上有策略性取捨，需進一步探討。另外 *late ratio* 雖無顯著，但  $\log\text{-odds}$  為負值符合一般預期。

點擊行為中，*input mean clicks* 為顯著負向預測因子 ( $\log\text{-odds} = -3.96, p = .007$ )，而 *input click cv* 以及 *output click cv* 則為正向因子，顯示點擊活躍程度波動高可能有助於學習成效，其他與間隔相關的點擊行為則皆無明顯效果。

此外，模型的隨機效應部分亦有所變化。Full Model 的群內相關係數 (ICC) 從 0.09 提升至 0.29，代表課程模組對學生表現的群體差異在控制行為變項後更加明顯。

模組的隨機變異 ( $\tau_{00}$ ) 也由 0.32 提升至 1.35，說明模型更敏感於課程本身所帶來的異質性。



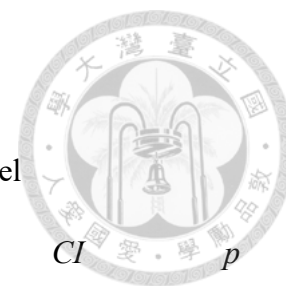
表 1

混合效應模型分析結果

<i>Predictors</i>		Base Model			Full Model		
		<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>
(Intercept)		-0.33	-1.04 – 0.39	0.368	-12.56	-14.60 – 10.53	< .001
region	East Midlands Region	0.02	-0.36 – 0.40	0.915	0.17	-0.43 – 0.77	.584
	Ireland	-0.30	-0.93 – 0.33	0.353	-0.17	-1.16 – 0.82	.738
	London Region	0.36	0.02 – 0.70	<b>0.040</b>	0.44	-0.11 – 1.00	.118
	North Region	-0.39	-0.90 – 0.11	0.129	-0.07	-0.87 – 0.72	.855
	North Western Region	-0.33	-0.71 – 0.04	0.080	-0.55	-1.15 – 0.04	.066
	Scotland	-0.04	-0.36 – 0.29	0.818	0.10	-0.41 – 0.62	.694
	South East Region	-0.21	-0.60 – 0.17	0.272	-0.61	-1.21 – -0.01	<b>.045</b>
	South Region	-0.13	-0.47 – 0.21	0.456	-0.22	-0.76 – 0.31	.412
	South West Region	0.00	-0.35 – 0.35	0.986	0.07	-0.49 – 0.63	.799
	Wales	-0.14	-0.54 – 0.27	0.518	-0.45	-1.11 – 0.20	.176

表 1 (續)

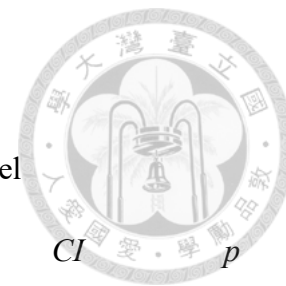
混合效應模型分析結果



<i>Predictors</i>	Base Model			Full Model			
	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	
West Midlands Region	-0.76	-1.14 – -0.37	< <b>0.001</b>	-1.25	-1.84 – -0.66	< <b>.001</b>	
Yorkshire Region	-0.11	-0.52 – 0.30	0.586	-0.38	-1.00 – 0.24	.232	
highest education	HE Qualification	0.38	0.17 – 0.60	< <b>0.001</b>	0.08	-0.26 – 0.42	.643
	Lower Than A Level	-0.71	-0.91 – -0.51	< <b>0.001</b>	-0.62	-0.93 – -0.31	< <b>.001</b>
	No Formal quals	-0.12	-1.38 – 1.14	0.854	0.56	-1.96 – 3.09	.662
	Post Graduate Qualification	1.38	0.54 – 2.22	<b>0.001</b>	0.69	-0.64 – 2.02	.310
age_band	35-55	0.09	-0.09 – 0.27	0.333	-0.30	-0.59 – 0.00	.050
	55<=	0.98	0.24 – 1.72	<b>0.010</b>	0.20	-1.04 – 1.43	.753
imd_band	10-20%	-0.02	-0.44 – 0.41	0.941	0.09	-0.59 – 0.77	.791
	20-30%	0.41	0.01 – 0.80	<b>0.046</b>	0.33	-0.30 – 0.95	.305
	30-40%	0.01	-0.38 – 0.40	0.951	0.04	-0.59 – 0.67	.908
	40-50%	0.08	-0.32 – 0.48	0.690	0.03	-0.61 – 0.67	.928

表 1 (續)

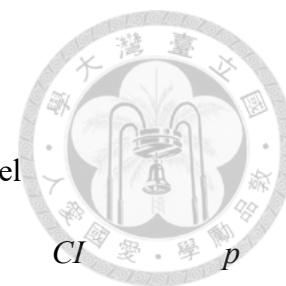
混合效應模型分析結果



<i>Predictors</i>	Base Model			Full Model			
	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	
50-60%	0.17	-0.23 – 0.57	0.408	0.12	-0.53 – 0.77	.711	
60-70%	0.19	-0.21 – 0.59	0.348	0.20	-0.44 – 0.83	.544	
70-80%	0.37	-0.02 – 0.76	0.065	0.08	-0.54 – 0.71	.792	
80-90%	0.36	-0.03 – 0.75	0.071	0.00	-0.62 – 0.62	.993	
90-100%	0.47	0.08 – 0.87	<b>0.019</b>	0.29	-0.33 – 0.92	.356	
gender	M	0.17	0.00 – 0.34	<b>0.047</b>	0.42	0.15 – 0.69	<b>.002</b>
date registration		0.08	-0.48 – 0.65	0.769	0.31	-0.58 – 1.20	.494
navigation	mean clicks			3.28	-0.64 – 7.20	.101	
	mean click interval			-2.14	-7.06 – 2.78	.394	
	click cv			-2.53	-6.43 – 1.38	.205	
input	mean clicks			-3.96	-6.83 – -1.09	<b>.007</b>	
	mean click interval			1.07	-1.99 – 4.13	.492	
	click cv			2.55	0.69 – 4.41	<b>.007</b>	
output	mean clicks			2.18	-0.26 – 4.61	.080	
	mean click interval			-1.68	-7.03 – 3.67	.539	
	click cv			2.13	0.01 – 4.25	<b>.049</b>	

表 1 (續)

混合效應模型分析結果



<i>Predictors</i>	Base Model			Full Model		
	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>	<i>Log-Odds</i>	<i>CI</i>	<i>p</i>
early score				5.23	3.95 – 6.50	<.001
mid score				7.48	6.17 – 8.79	<.001
late score				4.54	3.66 – 5.42	<.001
late ratio				-0.55	-1.47 – 0.36	.235
completion ratio				-0.81	-1.55 – -0.06	.035
<b>Random Effects</b>						
$\sigma^2$	3.29			3.29		
$\tau_{00}$	0.32	module_presentation		1.35	module_presentation	
ICC	0.09			0.29		
N	6	module_presentation		6	module_presentation	
Observations	2819			2819		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.091 / 0.173			0.713 / 0.796		

整體而言，混合效應模型分析結果顯示，雖然基本人口變項對學生成效具有些微解釋力，但需結合行為與階段性表現等衍生變項，方能有效提升模型預測能力與解釋性。尤其在 Full Model 中，不同類型的點擊行為與作業得分，對學習結果具有顯著區辨性。

### 第三節 類神經網路模型



#### 壹、SHAP 分析圖解讀方式

為進一步了解學生背景變項和學習成效間的關聯性，並藉此分析針對深度模型訓練結果施行合理性檢驗（sanity check），本研究透過 SHAP（SHapley Additive exPlanations）方法計算並繪製變項重要性總結圖，該圖能有效協助分析者理解模型的決策依據，以下將簡單說明總結圖之閱讀方式，以利讀者後續理解。

首先，每個 SHAP 值是針對「每一筆樣本」與「每一個變項」所計算的，圖中的每一個點代表一筆 SHAP 值，也就是該變項對該樣本預測結果所造成的貢獻量。圖中 X 軸為 SHAP 值，其數值可正可負，表示該變項對模型預測機率之變化幅度的影響力。若 SHAP 值為正，代表該變項的值促使模型更傾向預測為「高分組」；若為負，則表示該變項使模型傾向預測為「低分組」。

圖像左側會依序列出模型中最具影響力的前二十個變項，依據其平均 SHAP 值的絕對值排序，由上至下排列，表示對模型輸出貢獻的相對重要性，排越上方代表該變項對預測結果的影響越大。值得注意的是，圖中左側所列的名稱不全然對應於原始變項本身，而可能為類別變項經過 one-hot 編碼後所展開的個別類別。例如，變項 highest\_education 經編碼後會產生多個以其類別值為名稱的新欄位，如 highest\_education\_Lower Than A Level、highest\_education\_HE Qualification 等。SHAP 值在此情境下所代表的，是該類別在該樣本中出現（值為 1）時，對模型預測結果所造成的影響。因此，閱讀此圖時須注意，若多個欄位名稱來自同一原始變項，可綜合解讀其整體影響力，而非將其視為獨立變項。

圖中每個點的顏色則表示該樣本在該變項上的實際取值高低，紅色代表變項取值較高，藍色則代表取值較低；若為類別變項，則紅色代表該類別在樣本中出現，藍色意指未出現。綜合觀察顏色及 SHAP 值有助於理解變項值的變化方向是否與預測傾向一致，例如，若一個變項在高值時的 SHAP 值多數為正，則可推論該變項的高值傾向與高分表現相關。

以下將接續描述兩種類神經網路模型的分析結果。



## 貳、Base Model

### 一、訓練結果

為評估類神經網路模型在僅使用基本人口統計變項 (Base Model) 條件下的分類能力，本研究採用 5 折分層交叉驗證 (5-Fold Stratified Cross Validation) 進行模型效能測試。每一折交叉驗證皆重新執行完整的資料前處理與模型訓練流程，並以準確率 (Accuracy) 作為主要評估指標。模型於五折交叉驗證中所得準確率如下分別為 0.63、0.62、0.62、0.63、0.64，其平均準確率為 0.63，標準差為 0.0095，整體表現穩定。

此結果顯示，在未納入行為與表現等衍生變項的條件下，類神經網路模型仍具些微預測力，能辨識出部分與學習成效顯著相關之人口變項。然而，由於該模型僅使用學生背景資料，其準確率尚未達到高度辨識水準，亦未展現顯著分類能力，顯示此基準模型仍有進一步特徵擴充的空間。

### 二、SHAP 值分析

Base Model 之 SHAP 值分析結果圖中顯示，對於模型預測學生是否屬於高分組，最具影響力的背景變項之一為 highest education。其中又以類別 Lower Than A Level 有最顯著的效果，此變項為強烈的負向貢獻因子，代表若學生最高教育程度低於 A Level，則其被預測為高分組的機率顯著降低。從圖中可見到該列所有紅色樣本點 (意即所有 highest education 變項類別為 Lower Than A Level 的樣本) 都落在 SHAP 值小於零的區間，顯示此類別影響相當穩定。而相同變項中的類別 HE Qualification (大學及同等教育程度) 結果顯示較高學歷背景對預測高分有正向貢獻，顯示教育程度為學習表現的關鍵影響因子之一，此結果符合一般預期。

多個課程模組 (module\_presentation) 在模型中具有一定區辨力，顯示不同課程及開課時間本身即造成學生表現差異。其中部分模組 (如 DDD2013B) 對高分預測有顯著負面影響，可能與課程難度、學生組成或教材設計等因素相關。另外變項 imd\_band\_90-100% (代表較高社經地位) 及 region London Region (居住在倫敦地區) 也對模型預測結果為高分組的機率有正向影響，這些結果亦符合一般預期。

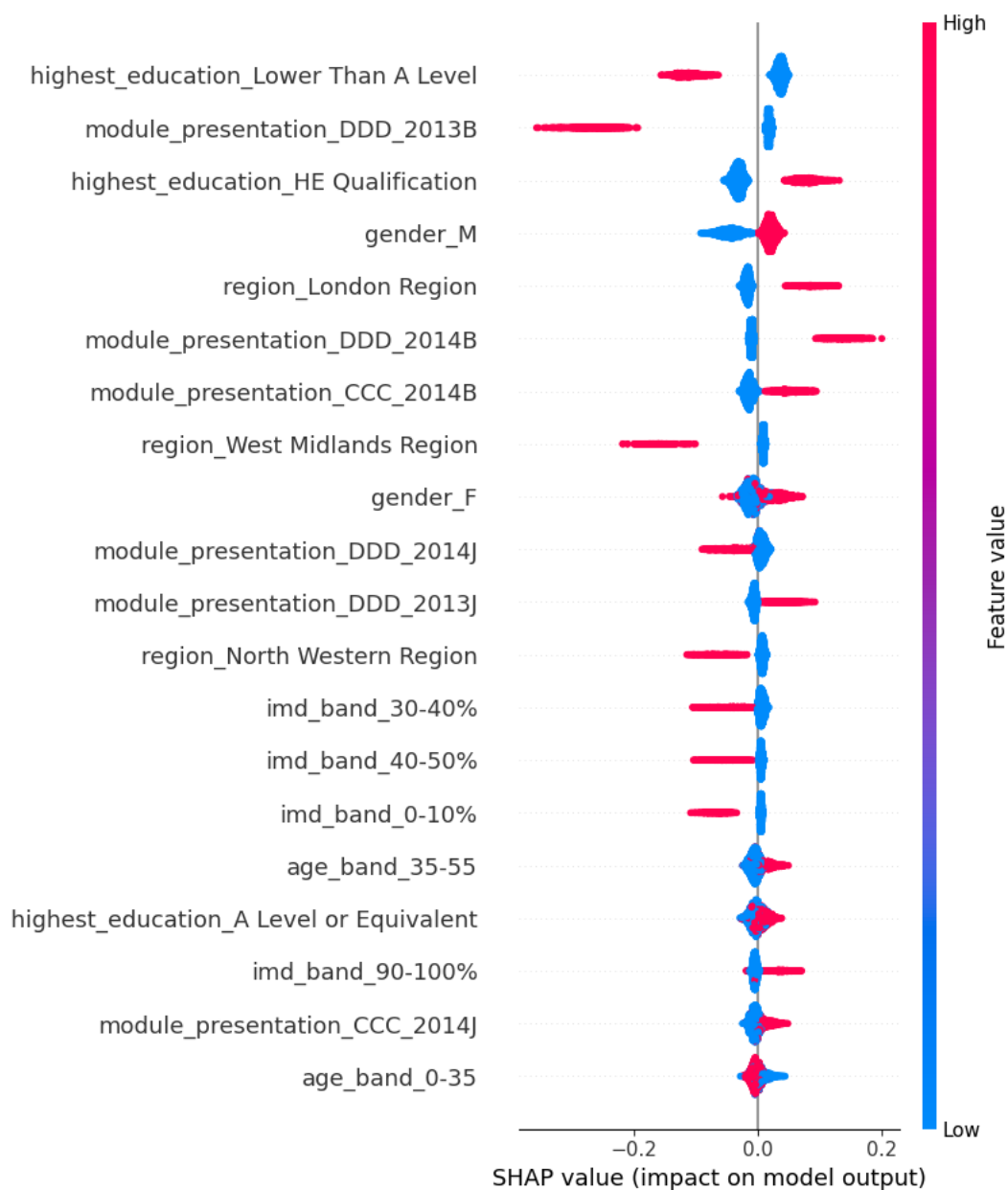
值得注意的是，模型本身正確率偏低，且許多特徵的 SHAP 值分布仍接近 0，此

分析僅作為後續與 full model 比較、解釋行為變項貢獻之基準。



圖 2

Base Model SHAP 值分析結果圖





## 參、Full Model

### 一、訓練結果

在納入學生學習行為與階段性表現等衍生變項後，本研究建立類神經網路之 Full Model，並同樣採用 5 折分層交叉驗證 (5-Fold Stratified Cross Validation) 評估其分類效能，並以準確率作為主要評估指標。五次交叉驗證所得準確率分別為 0.87、0.88、0.87、0.86、0.88，其平均準確率為 0.87，標準差為 0.0072，整體表現穩定且顯著優於 Base Model 的結果 (0.63)。

此結果顯示，將學習過程中可觀測的動態行為特徵與作業表現納入模型後，可大幅提升預測學生期末學習成果的準確性，且模型在不同資料切分下之效能維持高度一致，具備良好的泛化能力。

### 二、SHAP 值分析

為進一步解釋類神經網路 Full Model 中各特徵對預測結果的貢獻方向與重要性，同樣應用 SHAP 方法呈現模型預測邏輯的可解釋性結果。

由圖可見，與 Base Model 相比較，引入學習行為與作業表現相關的衍生變項後，特徵重要性排序產生明顯變化，且模型預測更顯著依賴學生的階段性表現。

在所有特徵中，mid\_score (課程中期的平均評量成績) 為最具貢獻力的預測變項，其 SHAP 值分布明確且主要集中於正向，顯示中期表現越高的學生，其被預測為高分組的機率也顯著提升。其次為 late\_score 與 early\_score，顯示學生在課程各階段的作業表現皆對學習成果具高度預測力，且三者影響方向一致，整體貢獻量大，反映評量成績為模型主要參考依據之一。

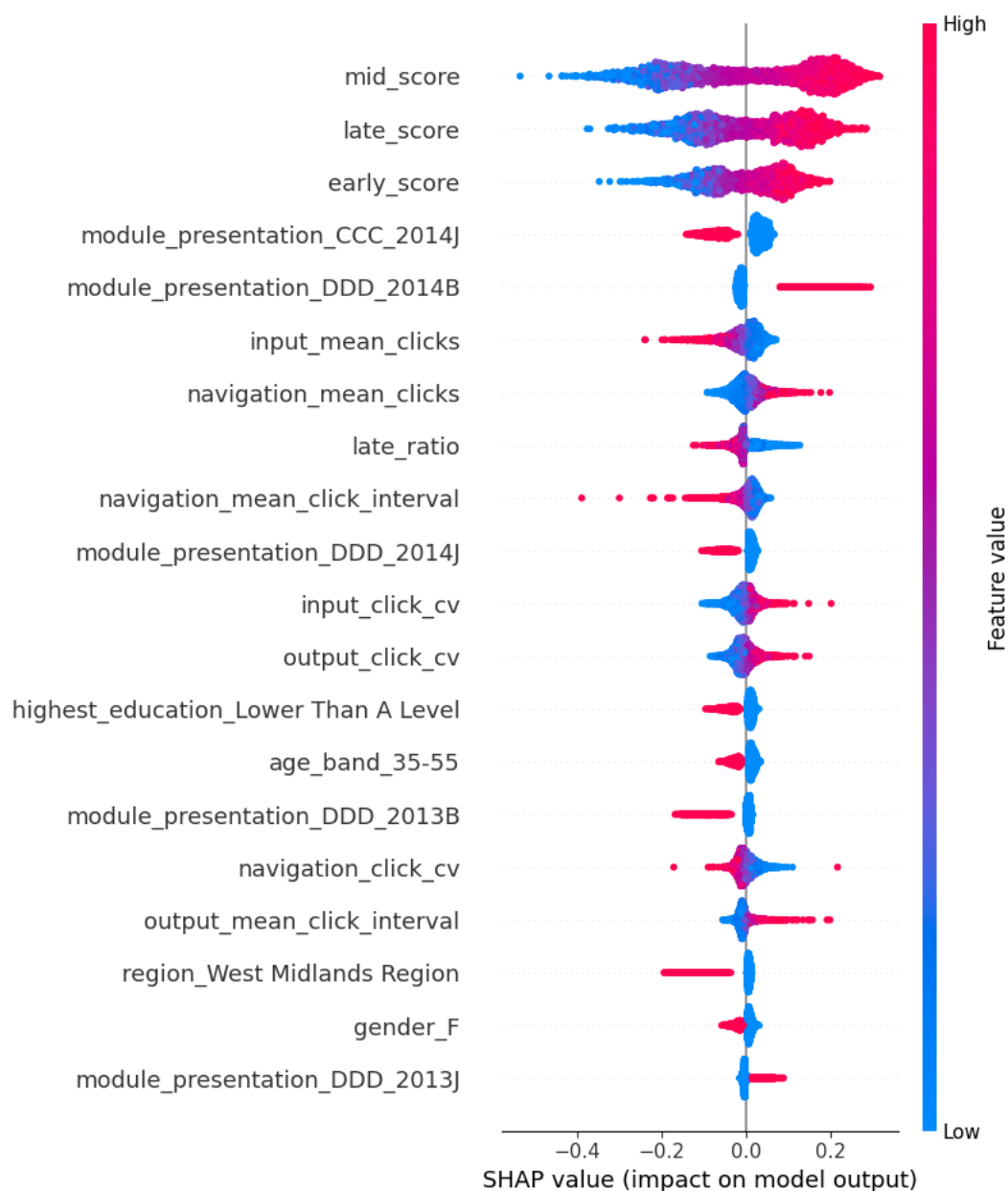
此外，late\_ratio (遲交比率) 亦位居前段位置，其 SHAP 值主要為負，顯示學生若有較高比例作業逾期，較可能被預測為低分組，呼應實務中時間管理能力與學習成效之關聯。

在學習行為相關的變項中，數個與點擊模式有關的特徵呈現出對預測結果具有代表性的貢獻，且其 SHAP 值分布具方向性，說明這些行為變項不僅具預測力，亦可揭示潛在的學習策略與行為型態。首先，navigation\_mean\_click\_interval (導覽型資源的

點擊間隔時間平均值) 呈現穩定負向 SHAP 值，表示若學生在瀏覽課程首頁、導覽資源的行為間隔越大，則越傾向被預測為低分組。navigation\_click\_cv 同為負向貢獻，變異係數本身反映學生的點擊數波動程度，顯示波動大、不規律的瀏覽行為模式傾向被預測為低分組，其意涵可能與學習策略行為相關。navigation\_mean\_clicks 則展現正向貢獻，表示瀏覽行為的總量高對學習成效可能有正向幫助。

圖 3

Full Model 之 SHAP 值分析結果圖



input\_mean\_clicks (輸入型資源的平均點擊次數) 整體呈現負向 SHAP 值分布，且變項值越高的樣本對模型預測高分組的貢獻反而越低，說明過度頻繁地點擊輸入型資源可能是成效有限的學習行為。而 input\_click\_cv (互動點擊次數的波動程度) 則呈現正向貢獻，顯示若學生使用互動資源時點擊模式為不規律、高度波動，反而與高分組的預測呈正相關。此一現象可能意味著高成效學生傾向在輸入互動資源間進行更多思考、內化與準備，並會選擇性地投入時間，顯示更具計畫性的行為模式。換言之，點擊多不必然代表投入深，反而可能指涉重複性高的無效行為。

最後，output\_click\_cv 和 output\_mean\_click\_interval 同為正向因子，說明知識運用行為中適當的間隔行為可能有助於學習的提升，並且高度波動的行為模式也提供正向影響。

綜合上述，這些點擊行為特徵揭示：學習成效較高的學生並非盲目頻繁操作，而是呈現有所調整、策略性明確的互動行為，並且不同類型的學習行為會展現不同模式；而學習成效偏低者則可能表現為點擊節奏單調、次數過高。這些細部行為特徵不僅提升模型預測能力，亦可作為實務上早期識別風險學生的重要線索，輔助模型辨識學生在學習過程中的投入模式與節奏。

值得注意的是，Base Model 中具高度影響力的靜態人口變項 (如 highest\_education\_Lower Than A Level、region\_West Midlands Region、gender\_F 等) 在 Full Model 中仍保有可見的 SHAP 值，但其貢獻明顯下降，排序也後移，反映這些背景資訊雖具有一定預測意義，但一旦引入學習行為與表現資料，其相對影響即被更具代表性的動態變項所取代。

整體而言，Full Model 的 SHAP 分析顯示：模型預測已從以背景條件為主，轉向依賴學習歷程中的具體行為與表現，模型決策邏輯更加聚焦於學生在課程期間的真實投入與成果，這亦說明引入衍生變項對模型可解釋性與預測力的雙重增益。



## 第四章 綜合討論

本章共分為三節，第一節討論本研究結果帶來的新發現與貢獻，第二節則提出研究方法的限制與不足，最後一節講述此研究主題可能的未來發展方向，以供後續研究參考。

### 第一節 研究發現與貢獻

#### 壹、對教學者的啟示：遠距教學中的早期干預

本研究結果顯示，學生無論是前期、中期或學期末的學習表現皆對最終學習成效展現出相當的影響力，顯示任一階段成績並非唯一的關鍵指標。綜觀學生於學期三階段中的作業得分表現，中期成績為各模型中預測力最強的變項，初期成績亦呈現穩定且正向的預測效果，雖然後期成績同樣達到統計顯著，而其貢獻程度並未顯著優於其他階段。

本研究的結果不僅指出前期學習表現對最終成果同樣具有穩定影響力，也反映出線上課程中潛藏的風險。通常教師可能預期學生對課程內容的掌握會隨時間累積而逐漸提升，因此容易將初期的表現不佳視為正常現象，同時在遠距教學的情境下，授課者無法如實體課堂般即時觀察學生的反應與學習狀態。如此一來便導致在學期初即陷入困難的學生更難被及時發現與協助，這種教學感知的不足更可能使學生的學習落後在無人察覺的情況下逐漸擴大，而未能隨課程進展而改善。因而「早期干預」的重要性尤顯迫切，過去研究亦提及遠距教學中干預設計的重要性並試圖提出可能的預測方法(Figueroa-Cañas & Sancho-Vinuesa, 2020; Z. Wang 等, 2024; Azimi 等, 2020)。若能於早期階段辨識學習表現落後者，教育者便可及早啟動補救措施，例如提供即時回饋、

加強學習策略引導或安排額外學習資源，對於提升整體學習成效與降低後期風險具有實質意義。

由此可見，線上教學的普及雖為學習提供了更多彈性與資源，卻也同時需要及時的干預機制來彌補授課者與學生間的資訊落差。幸而，線上學習平台本身即具備易於觀測與追蹤的技術特性，能夠即時蒐集並回報學習者的評量表現，為早期預警與個別化干預創造可能。

## 貳、遠距學習中策略規劃行為的重要性

用以描述評量繳交情形的遲交比率與作業完成率皆在模型中呈現負向貢獻，顯示學生的缺交行為可能是一種學習策略。首先，作業完成度越高並不保證越好的學習成效，與過去的研究結果有差異(Cooper 等, 2006; Morris 等, 2005; Wu 等, 2024)。然而與此同時，遲交率高卻仍然預示更差的學習表現，此處需注意遲交比率指的是學生有繳交的評量中遲交的比率，並不包含未繳交的評量。過去研究顯示，過長的上課時數容易導致作業品質的下降，進而影響學習成效，而過多的作業也可能增加學生的認知負荷及心理疲累程度，可能降低學習動力並減弱學習成效(Chin 等, 2022; Guo 等, 2024; Núñez 等, 2015)。此結果顯示高分群學生可能受益於線上學習平台高度的自主性，能夠根據自身狀況計劃性地選擇欲完成的評量，適度調節評量帶來的學習壓力。此外，考慮資料集中準時提交的行為仍然能體現學生對該評量的認真參與，說明其缺交行為可能並非源於對課程參與度低或較差的時間管理能力導致，而是高層次的時間管理與任務選擇能力。然而，單看完成率無法完全瞭解學生繳交的評量品質，以及評量設計是否能真實反映學生的學習狀況，此處仍需由未來研究補足。

而就點擊行為相關的分析結果顯示，知識獲取行為平均點擊次數為負向預測因子，說明單純的高頻率操作並不同於有效投入，符合過去研究的發現 (Hadwin 等, 2007; Misanchuk & Schwier, 1992)。學習成效較高的學生不一定會進行頻繁的操作，反而通常表現為選擇性強且學習節奏清晰。知識獲取行為及應用行為的平均間隔天數都對學習成效有正向影響，表明無論是知識的吸收或運用，都能透過安插適度的休息間隔來提升最終效益。然而，資源瀏覽行為之平均間隔天數卻呈現負向影響，此一發現與間隔效應的理論有所衝突。瀏覽行為發生在學生於不同資源頁面之間切換時，屬於一種跨越知識獲取及應用的規劃行為，其本身與實質上的學習行為關聯較弱，而是屬

於「學習策略規劃」的範疇。同時，資源瀏覽行為之波動程度亦呈現負向影響，本研究認為規律的瀏覽行為可能有助於學習策略的規劃，較短的點擊間隔配合穩定的點擊次數，可能顯現高分組學生在修課期間有持續建立學習策略的傾向。

然過去較少研究提及資源瀏覽、規劃相關的行為表現，此行為分類與學習成效的實際關聯仍待更多研究證實，因此，本研究期能以此結果提起研究者對線上學習中「策略規劃」行為的重視。

### 參、線上學習行為的分析方法

本研究提出一種新的線上學習行為分析方法，強調行為解釋不應僅僅依賴「多與少」的頻率比較，而應轉向「策略與選擇」的層面。特別是在線上學習環境中，學生對自身學習行為有更高的掌控度與彈性，這使得學習行為的多樣性與策略性成為影響學習成效的重要因素。簡單的行為頻率或次數分析往往無法充分揭示學習行為的實質意圖與效果，因為不同類型的學習行為在學習過程中的作用和影響可能大不相同。

此外，本研究進一步強調了學習行為分類的重要性。透過將學習行為細分為「資源瀏覽行為」（Navigation）、「知識獲取行為」（Input）和「知識應用行為」（Output），我們發現這些行為的間隔效應對學業成效的影響有所不同。例如，知識獲取行為的平均間隔天數與學業成效呈現正向關聯，而資源瀏覽行為的間隔過長則有負向預測作用。這表明，單純將這些行為混合在一起進行分析，可能會因為不同類型行為間的效果相互抵銷，從而無法揭示其潛在的差異性。因此，分類後進行單獨分析能夠更準確地捕捉到行為間的微妙差異，並為學習成效提供更具解釋力的預測因子。

總結來說，本研究提供了一個有效的框架，將學習行為進行分類並深入分析其不同間隔效應，突破了傳統行為分析的限制，為未來的學習行為研究提供了更多元的分析視角。

### 第二節 研究限制

本研究透過結合混合效應模型與類神經網路方法，並輔以 SHAP 解釋技術，成功建構具預測力與可解釋性的學習成效模型，且揭示多項具代表性的學習行為特徵。然而整體研究仍存在若干限制，將於此節加以說明，並提出未來可能的改進方向，俾使

後續研究與應用能有更準確的判讀與合理的延伸使用。



## 壹、資料來源與結構性限制

本研究所使用之 OULAD 資料集為英國著名的 Open University 所提供的學習分析開放資料集，具備一定規模與品質，然其本質上屬於次級資料（secondary data），資料收集方法與結構並非專為本研究目的設計。首先，資料涵蓋時間為 2013–2014 年，平台與課程設計可能與當前線上學習環境有所落差，此外學生的學習需求也可能隨時代改變而有變異，故研究結果需謹慎遷移至今日教育場域。再者，資料集中僅部分課程設有期末考試成績，導致本研究能納入之樣本受限，可能對模型結果造成潛在影響。

另外，OULAD 所記錄之學習行為主要為 VLE 點擊資料，這類點擊紀錄雖提供操作層次上的客觀數據，卻難以反映學生背後的學習動機、認知投入或實際理解程度。舉例而言，點擊某資源次數無法反映實際停留在該資源的時間，或學習過程中的深度處理程度。同時，OULAD 以日作為結算單位，因此未能詳細捕捉學生在單日內的學習行為模式，而可能遺失了潛在的有效資訊。

針對此問題，未來可考慮將點擊行為的結算時間縮短，引入更細緻的時間相關變項（如實際停留時間、任務間切換頻率），並結合平台紀錄中可取得的其他訊息，如討論區留言內容、影片播放狀態或速度等，以建立更全面的學習行為輪廓。

## 貳、方法選擇與模型詮釋的限制

在方法層面，混合效應模型與類神經網路雖各有優勢，但亦帶來不同程度的限制。混合效應模型的主要限制在於對高維交互作用與非線性結構的處理能力有限，儘管其結果具可解釋性，但對於行為變項與學習成效間潛在的非線性關聯與複雜結構可能無法充分掌握。相對地，類神經網路模型雖具備強大擬合能力，且 SHAP 解釋技術有效提升其透明度，但其解釋結果仍需謹慎處理。例如，SHAP 為後設分析，其所揭示的機制並非模型決策過程中的直接邏輯，而為對結果的近似推論。

此外，本研究所使用之特徵工程（如分類計算點擊變異係數、學習間隔等）雖經理論與經驗設計，但仍為人為定義變項，未必能涵蓋所有可能的學習行為類型。

總體而言，以上限制並不削弱本研究所提出的發現與分析價值，惟在實務應用與未來研究設計時，應充分理解資料與方法背後的假設與適用範圍，方能做出更精準且具解釋力的教育決策與學術貢獻。



### 第三節 未來展望

#### 壹、建立早期干預機制

本研究顯示，學生在學期初期與中期的學習表現皆對期末成績具有顯著的預測能力，這一發現強調了早期干預在遠距教學中的重要性。未來的研究可以進一步探討如何通過早期的學習表現進行精確的預測，並進行針對性的干預措施。可以結合本研究提出的更多元的學習資訊，如經分類的學習行為頻率、時間間隔，以進一步提高早期預測的準確性。這些模型可幫助教師更及時、有效地辨識出學習上的問題，從而進行針對性的支持與幫助。

未來亦可以探索不同類型的早期干預策略在遠距教學中的效果，進一步探討哪些具體的干預措施能在不同的學習階段及情境中發揮最佳效用。例如，是否不同的學生群體（如教育背景不同的族群）適用不同的干預策略？不同的學科內容或學習形式是否會影響干預的效果？這些問題都值得進一步研究。

此外，如何利用線上學習平台的技術優勢來實現個別化學習支持，將是未來研究的重點。平台可透過預測模型標記學期初出現不穩定行為的學生，自動發出提醒或建議，亦可依據學生的表現分群，將程度相近的學習者引導至差異化的學習資源或課程軌道，將這些學習數據與教師的教育判斷相結合，並根據每位學生的需求及時調整教學策略，如針對學習動機不高或學習進度較慢的學生，設計更具吸引力的干預機制

總體而言，線上學習環境下雖缺乏傳統教室的即時互動與察覺能力，卻同時具備前所未有的資料可視性與資訊技術支援，只要善加研究與設計，早期干預機制不僅可行，且有潛力比傳統課堂更精準與有效。

## 貳、深入探討策略規劃行為



本研究揭示了學習規劃行為在學習過程中的重要性，尤其是在資源瀏覽與策略規劃方面，未來的研究可以進一步詳細探討不同學習規劃行為對學習成效的影響。首先，學習規劃行為與學習成效之間的關聯仍有待進一步驗證。未來研究可以設計長期追蹤研究，深入探討學習規劃行為的長期影響，特別是在學習者面臨複雜課程或跨領域學習任務時。另外也可以從學習者的個人特徵出發，探討不同學習風格、動機或學習策略偏好對學習規劃行為的影響。例如，學習者的時間管理能力、目標設定方式或自我監控能力，可能會顯著改變他們在學習過程中的規劃行為與資源利用方式。透過這些變數的深入分析，研究者能更全面地理解學習策略，提升學習規劃行為的有效性。

其次，未來研究可以探討學習規劃行為在不同學習平台或學習模式下的表現。隨著線上學習與混合式學習日益普及，不同的學習環境可能會對學習者的規劃行為產生不同的影響。具體來說，學習平台的界面設計、資源呈現方式、互動功能等因素可能會影響學習者如何進行學習規劃。因此，未來的研究可以比較不同平台間學習規劃行為的差異，並提出如何優化學習平台設計以促進學習規劃的具體建議。

最後，學習規劃行為的發展和訓練也是未來研究的一個重要方向。儘管學習規劃行為對學習成效有顯著影響，但許多學習者可能未能有效掌握適當的規劃方法。未來的研究可以探索如何通過干預措施，如學習規劃指導、時間管理技巧培訓等，來幫助學習者提升其規劃能力。這類研究可以評估不同干預方式對學習成效的長期影響，並揭示哪些策略最能幫助學習者在自主學習環境中發揮最佳效能。

## 第四節 研究結論

本研究旨在深入探討影響線上學習成效的關鍵指標，特別是學習行為與學業表現之間的關聯。透過分析學生線上互動行為，包含作業繳交情形，並將學習行為分為「資源瀏覽行為」（Navigation）、「知識獲取行為」（Input）與「知識應用行為」（Output）的點擊三類，分別探討次數、行為分布及休息間隔。研究發現學習行為的節奏對學業成效有著顯著的預測能力，高成效學生通常展現出更加策略性的學習行

為，例如適當的學習間隔與選擇性的作業繳交。

本研究不僅揭示了學習過程中策略規劃行為的重要性，區分適用及不適用間隔效應的學習行為種類，也分析學生階段性表現的預測力，提出早期成績對最終表現的高度影響，藉此強調了在遠距學習環境中早期干預的必要性。

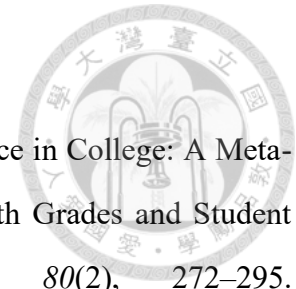




## 參考文獻

- Azimi, S., Popa, C.-G., & Cucić, T. (2020). Improving Students Performance in Small-Scale Online Courses—A Machine Learning-Based Intervention. *International Journal of Learning Analytics and Artificial Intelligence for Education (iJAI)*, 2(2), Article 2. <https://doi.org/10.3991/ijai.v2i2.19371>
- Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of Foreign Language Vocabulary and the Spacing Effect. *Psychological Science*, 4(5), 316–321. <https://doi.org/10.1111/j.1467-9280.1993.tb00571.x>
- Bélanger, P., & Tuijnman, A. (1997). *New patterns of adult learning: A six-country comparative study* (1st ed). Pergamon ; Unesco Institute for Education.
- Carvalho, P. F., Sana, F., & Yan, V. X. (2020). Self-regulated spacing in a massive open online course is related to better learning. *Npj Science of Learning*, 5(1), 2. <https://doi.org/10.1038/s41539-020-0061-1>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>
- Chin, J. M.-C., Lin, H.-C., & Chen, C.-W. (2022). Homework and learning achievements: How much homework is enough? *Educational Studies*, 48(3), 408–423. <https://doi.org/10.1080/03055698.2020.1766423>
- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does Homework Improve Academic Achievement? A Synthesis of Research, 1987–2003. *Review of Educational Research*,

76(1), 1–62. <https://doi.org/10.3102/00346543076001001>



Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class Attendance in College: A Meta-Analytic Review of the Relationship of Class Attendance With Grades and Student Characteristics. *Review of Educational Research*, 80(2), 272–295. <https://doi.org/10.3102/0034654310362998>

Eren, O., & Henderson, D. J. (2006). The Impact of Homework on Student Achievement. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.917447>

Eren, O., & Henderson, D. J. (2011). Are we wasting our children's time by giving them more homework? *Economics of Education Review*, 30(5), 950–961. <https://doi.org/10.1016/j.econedurev.2011.03.011>

Figueroa-Cañas, J., & Sancho-Vinuesa, T. (2020). Early Prediction of Dropout and Final Exam Performance in an Online Statistics Course. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 15(2), 86–94. <https://doi.org/10.1109/RITA.2020.2987727>

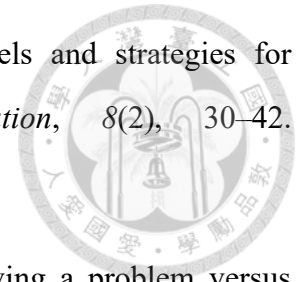
Grodner, A., & Rupp, N. G. (2011). *The Role of Homework in Student Learning Outcomes: Evidence from a Field Experiment* (SSRN Scholarly Paper 1892173). Social Science Research Network. <https://doi.org/10.2139/ssrn.1892173>

Guo, L., Li, J., Xu, Z., Hu, X., Liu, C., Xing, X., Li, X., White, H., & Yang, K. (2024). The relationship between homework time and academic performance among K-12: A systematic review. *Campbell Systematic Reviews*, 20(3), e1431. <https://doi.org/10.1002/cl2.1431>

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2), 107–124. <https://doi.org/10.1007/s11409-007-9016-7>

Hillman, D. C. A., Willis, D. J., & Gunawardena, C. N. (1994). Learner-interface interaction

in distance education: An extension of contemporary models and strategies for practitioners. *American Journal of Distance Education*, 8(2), 30–42. <https://doi.org/10.1080/08923649409526853>



Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17(6), 649–667. [https://doi.org/10.1016/S0022-5371\(78\)90393-6](https://doi.org/10.1016/S0022-5371(78)90393-6)

Johnson, G. M. (2005). Student Alienation, Academic Achievement, and WebCT Use. *Educational Technology & Society*, 8(2), 179–189.

Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A., & Brooks, C. A. (2018). How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research*, 88, 43–86. <https://doi.org/10.3102/0034654317740335>

Kim, S. K., & Webb, S. (2022). The Effects of Spaced Practice on Second Language Learning: A Meta-Analysis. *Language Learning*, 72(1), 269–319. <https://doi.org/10.1111/lang.12479>

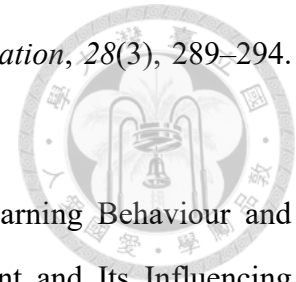
Kizilcec, R. F., & Halawa, S. (2015). Attrition and Achievement Gaps in Online Learning. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, 57–66. <https://doi.org/10.1145/2724660.2724680>

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, 104, 18–33. <https://doi.org/10.1016/j.compedu.2016.10.001>

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data*, 4(1), 170171. <https://doi.org/10.1038/sdata.2017.171>

Latif, E., & Miles, S. (2020). The Impact of Assignments and Quizzes on Exam Grades: A

Difference-in-Difference Approach. *Journal of Statistics Education*, 28(3), 289–294.  
<https://doi.org/10.1080/10691898.2020.1807429>



Li, J., & Xue, E. (2023). Dynamic Interaction between Student Learning Behaviour and Learning Environment: Meta-Analysis of Student Engagement and Its Influencing Factors. *Behavioral Sciences*, 13. <https://doi.org/10.3390/bs13010059>

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & Education*, 54(2), 588–599.  
<https://doi.org/10.1016/j.compedu.2009.09.008>

McDaniel, M. A., Fadler, C. L., & Pashler, H. (2013). Effects of spaced versus massed training in function learning: Journal of Experimental Psychology: Learning, Memory, and Cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(5), 1417–1432. <https://doi.org/10.1037/a0032184>

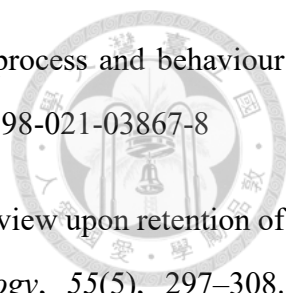
Misanchuk, E. R., & Schwier, R. A. (1992). Representing Interactive Multimedia and Hypermedia Audit Trails. *Journal of Educational Multimedia and Hypermedia*, 1(3), 355–372.

Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, 3(2), 1–7. <https://doi.org/10.1080/08923648909526659>

Morris, L. V., Finnegan, C., & Wu, S.-S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education*, 8(3), 221–231.  
<https://doi.org/10.1016/j.iheduc.2005.06.009>

Núñez, J. C., Suárez, N., Cerezo, R., González-Pienda, J., Rosário, P., Mourão, R., & Valle, A. (2015). Homework and academic achievement across Spanish Compulsory Education. *Educational Psychology*, 35(6), 726–746.  
<https://doi.org/10.1080/01443410.2013.817537>

Qiu, F., Zhang, G., Sheng, X., Jiang, L., Zhu, L., Xiang, Q., Jiang, B., & Chen, P.-K. (2022).

- 
- Predicting students' performance in e-learning using learning process and behaviour data. *Scientific Reports*, 12(1), 453. <https://doi.org/10.1038/s41598-021-03867-8>
- Reynolds, J. H., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology*, 55(5), 297–308. <https://doi.org/10.1037/h0040734>
- Schutte, G. M., Duhon, G. J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. Distributed practice on basic math fact fluency growth rates. *Journal of School Psychology*, 53(2), 149–159. <https://doi.org/10.1016/j.jsp.2014.12.003>
- Sun, Y. (2015). Characteristics of on-line learning behaviour in distance education. *China Educ. Technol*, 8, 64–71.
- Tawil, S., & Locatelli, R. (2015). *Rethinking education: Towards a global common good?* <https://unesdoc.unesco.org/ark:/48223/pf0000232555>
- Vlach, H. A., Kaul, M., Hosch, A., & Lazaroff, E. (2022). Attending less and forgetting more: Dynamics of simultaneous, massed, and spaced presentations in science concept learning. *Journal of Applied Research in Memory and Cognition*, 11(3), 361–373. <https://doi.org/10.1016/j.jarmac.2021.10.007>
- Wang, A. Y., & Newlin, M. H. (2002). Predictors of web-student performance: The role of self-efficacy and reasons for taking an on-line class. *Computers in Human Behavior*, 18(2), 151–163. [https://doi.org/10.1016/S0747-5632\(01\)00042-5](https://doi.org/10.1016/S0747-5632(01)00042-5)
- Wang, Y., Zuo, M., He, X., & Wang, Z. (2025). Exploring Students Online Learning Behavioral Engagement in University: Factors, Academic Performance and Their Relationship. *Behavioral Sciences*, 15(1), Article 1. <https://doi.org/10.3390/bs15010078>
- Wang, Z., Koprinska, I., & Jeffries, B. (2024). Interpretable Methods for Early Prediction

of Student Performance in Programming Courses. A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt, *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky* (115–123). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-64312-5\\_14](https://doi.org/10.1007/978-3-031-64312-5_14)

Wu, D., Li, H., Zhu, S., Yang, H. H., Bai, J., Zhao, J., & Yang, K. (2024). Primary students' online homework completion and learning achievement. *Interactive Learning Environments*, 32(8), 4469–4483. <https://doi.org/10.1080/10494820.2023.2201343>

Ye, M., Sheng, X., Lu, Y., Zhang, G., Chen, H., Jiang, B., Zou, S., & Dai, L. (2022). SA-FEM: Combined Feature Selection and Feature Fusion for Students' Performance Prediction. *Sensors*, 22(22), Article 22. <https://doi.org/10.3390/s22228838>

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education*, 29, 23–30. <https://doi.org/10.1016/j.iheduc.2015.11.003>

# 附錄



### 衍生變項的資料分布圖

