

國立臺灣大學醫學院職能治療學系碩士論文

School of Occupational Therapy

College of Medicine

National Taiwan University

Master Thesis



以機器學習模型發展及驗證短版中風衝擊量表

Using Machine Learning Algorithms to Develop and
Validate a Brief Version of Stroke Impact Scale

李沛綺

Pei-Chi Li

指導教授：謝清麟教授

Advisor: Dr. Ching-Lin Hsieh

中華民國 109 年 7 月

July 2020



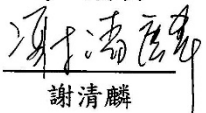
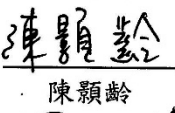
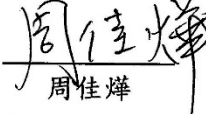
國立臺灣大學碩士學位論文

口試委員會審定書

以機器學習模型發展及驗證短版中風衝擊量表

Using Machine Learning Algorithms to Develop and Validate a Brief Version
of Stroke Impact Scale

本論文係李沛綺 (R07429008) 於國立臺灣大學職能治療學系所完成之
碩士學位論文，於民國 109 年 6 月 22 日經下列考試委員審查通過及口試及格，
特此證明。

考試委員	服務機構	職稱
 謝清麟	國立臺灣大學醫學院 職能治療學系所	教授
 陳穎齡	國立臺灣大學醫學院 職能治療學系所	副教授
 周佳燁	輔仁大學醫學院 職能治療學系	副教授

指導教授：謝清麟教授
職能治療學系系主任：薛漪平主任

中華民國 一百零九年 六 月

中文摘要



背景及目的：中風衝擊量表 (Stroke Impact Scale 3.0, SIS) 為著名之生活品質量表。然而，SIS 之題數過多 (共 59 題) 導致施測時間較長，不利於臨床常態評估。雖 SIS 已有多種短版測驗可縮短施測時間，但各短版 SIS 僅能提供單一分數代表個案之整體生活品質，尚無可提供各向度分數之版本，難以瞭解個案於各生活層面之情形。昔日研究發現機器學習可提供優於傳統線性迴歸模型之預測精準度，極具潛力可用於發展短版量表。本研究之目的有二：(1) 使用機器學習模型發展短版 SIS (Stroke Impact Scale-Brief Machine learning version, SIS-ML)；(2) 驗證其同時效度及收斂效度。

方法：本研究使用已發表論文之 SIS 資料進行模擬分析，分為二個階段：(一) 發展 SIS 短版；(二) 驗證其同時效度及收斂效度。發展 SIS-ML 分為 5 個步驟：(1) 挑選短版題目之題組：使用包含套鎖迴歸 (lasso regression) 之人工神經網路 (artificial neural network, ANN) 模型挑選 SIS 短版之題目組成總數 16 題至 32 題共 17 組 SIS-ML 題組。題目需包含 8 個向度，且各向度題目需包含 2 至 4 題；(2) 訓練 ANN 模型以預測各向度分數：為挑選預測力最佳之模型，研究者共訓練 136 種機器學習模型之預測力，包含：17 種題數組合 (16 至 32 題之版本)、2 種隱藏層數 (6 層及 10 層)，以及 4 種人工神經元個數 (每層 8、32、196、512 個神經元) 之模型。訓練資料以隨機順序並依據 35%、35% 及 30% 之人數比例，分為訓練資料、驗證資料及測試資料，以進行模型訓練、驗證及測試模型；(3) 以訓練資料挑選預測力較佳之 ANN 模型架構：挑選於各向度決定係數 (coefficient of determination) 超過 0.80 之模型架構；(4) 以驗證資料挑選預測力較佳之題數組合：自前一步驟之挑選結果中，各向度分數之決定係數超過 0.80 的題數組合 (即 SIS-ML 總題數)；(5) 以測試資料選擇最適之機器學習模型：自步驟 4 之結果中，挑選整體預測力最佳之模型，以做為 SIS-ML 分數計算之依據。步驟 3 與 4 之數據來自驗證資料，步驟 5 來自測試資料。若該題組有多個模型達標準，則取各組模型分數之平均，以選擇最適之題組及模型。於第二階段，研究者使用第一階段選定最適之模型，匯入測試資料以模擬原版 SIS 分數，並驗證模

擬分數與原版 SIS 分數、美國國家衛生研究院中風量表 (National Institutes of Health Stroke Scale, NIHSS)、巴氏量表 (Barthel Index, BI) 之相關程度。

結果：本研究使用之前研究收案之 256 位中風個案資料，並將資料分為訓練資料 89 人，驗證資料 90 人及測試資料 77 人。本研究於第一階段共 3 種模型架構 (6X196、6X512 及 10X196) 所衍生模型具有較佳之預測力 (步驟 3)；於此 3 種架構所衍生之模型中，共 6 種題數組合 (27 至 32 題) 之預測力較好 (步驟 4)。於前述之模型架構中，總題數 27 搭配 6X196 之模型架構之平均決定係數最高，故被選為最適之題組及模型架構之配對。本研究第二階段驗證 SIS-ML 之同時效度及收斂效度，同時效度驗證結果發現 SIS-ML 之模擬分數與原版 SIS 有高度相關 ($r = 0.92-0.99$)，此結果顯示 SIS-ML 與原版 SIS 有良好之同時效度。收斂效度驗證結果發現 SIS-ML 各向度之模擬分數與 NIHSS 呈低度至中度相關 ($r = -0.34 - -0.59$)，模擬分數與 BI 達低度至高度相關 ($r = 0.22 - 0.74$)。

結論：本研究結果支持機器學習模型發展之 SIS-ML 可將 SIS 之評估題目數量縮短至 45%，第一階段結果顯示最適之短版題數為 27 題。本研究第二階段亦支持 27 題之 SIS-ML 之模擬分數與原版 SIS 有良好之同時效度，與 NIHSS 及 BI 亦與原版相同之收斂效度。然而，本研究所使用之樣本數少，且 SIS-ML 尚未進行臨床驗證。因此未來研究需使用更大之樣本數確認目前版本及目前模型之預測力及再使用獨立之樣本驗證 SIS-ML 之信效度。

關鍵字：機器學習、中風衝擊量表、中風、短版、心理計量特性

英文摘要



Background and purposes: The Stroke Impact Scale (SIS) is a commonly used measure of health-related quality of life. However, the SIS has too many items that need much time to administer and place burden on patients. Although previous studies have developed brief versions, the brief versions cannot represent the original domains' scores. Researchers show that the machine learning can improve the accuracy of prediction. Thus, machine learning algorithms may assist with the development of a brief version of the SIS. The objectives of our study were: (1) using machine learning algorithms to develop a brief version of the SIS (SIS-ML); and (2) validating its concurrent validity and convergent validity in patients with stroke.

Methods: Our study used the data collected from a previous study for simulation analysis, and was comprised of 2 phases: (1) development of the SIS-ML; (2) validation of its concurrent validity and convergent validity. Phase 1 contained 5 steps to develop the SIS-ML: (1) choosing item groups of the SIS-ML: Two to four items in each domain were selected to form the SIS-ML using an artificial neural network (ANN) model with lasso regression. (2) training the ANN models to optimize the predictive power in the domain scores. A total of 136 models were trained, which were formed from 17 sets of items (number of items ranged from 16 to 32), 2 sizes of the hidden layers (6 and 10 layers), and 4 neurons in each layer (8, 32, 196, and 512 neurons). The data were separated randomly according to a ratio of 35%, 35%, 30% of the whole sample to become a training set, a validating set, and a testing set, respectively. (3) choosing the model frameworks with better predictive power by the training set. The models that achieved coefficients of determination (R^2) exceeding 0.80 were retained. (4) choosing the models with better predictive power by the validating set. The models that were retained in the previous step and achieved individual R^2 in each domain > 0.80 were retained. (5) choosing the best model with high predictive power and efficiency by the testing set. The model that used the fewest items to achieve the highest average R^2 was selected. The R^2 were calculated using the validating set and testing set for step 3 to step 5, respectively. In Phase 2, Pearson's correlation coefficient (r) was used to validate the concurrent validity between the SIS-ML, which was chosen in Phase 1, and the original

SIS, and the convergent validity between the SIS-ML and the National Institute of Health Stroke Scale (NIHSS) and Barthel Index (BI).

Results: In Phase 1, 17 item groups were chosen, resulting in 136 groups of model scores (steps 1 and 2). 3 models were chosen (6X196, 6X512, and 10X196) as the better model frameworks (step 3). 6 sets of items of the SIS-ML had better predictive power (step 4). 27 to 32 items were considered as the acceptable numbers of items. Finally, the best number of items was 27 and the best model framework was 6X196 (step 5). In Phase 2, the SIS-ML, which was chosen in Phase 1, had high correlation ($r = 0.92-0.99$) with the SIS. The SIS-ML had fair to medium correlation ($r = -0.34- -0.59$) in the NIHSS and had poor to high correlation ($r = 0.22 - 0.74$) with the BI.

Conclusion: The SIS-ML contain less than half of the items of the original SIS. The 27 items with the 6X196 model framework was the best version of the SIS-ML. The SIS-ML had good concurrent validity with the SIS and the convergent validity was similar to that of the SIS.

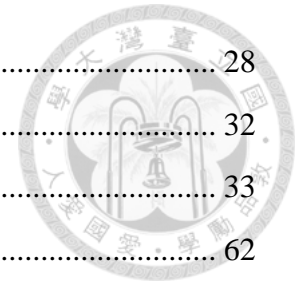
Keywords: machine learning, Stroke Impact Scale, stroke, short forms, psychometric property

目 錄



口試委員會審定書	i
中文摘要	ii
英文摘要	iv
圖目錄	viii
表目錄	ix
第一章 中風患者之生活品質簡介	1
第一節 中風之病理學簡介	1
第二節 健康生活品質簡介	3
第三節 評估中風患者之 HRQOL 之重要性	4
第二章 健康生活品質量表之介紹及評析	5
第一節 良好健康生活品質量表之主要特性	5
第二節 常用於中風患者之 HRQOL 量表	7
第三節 比較 HRQOL 量表及發展短版量表之選擇	9
第三章 運用機器學習模型發展短版量表	10
第一節 機器學習簡介	10
第二節 人工神經網路	12
第三節 機器學習之實作	15
第四節 運用「機器學習模型」發展短版測驗之優勢	16
第四章 研究目的	18
第五章 研究方法	19
第一節 資料來源及研究對象	19
第二節 第一階段：發展 SIS-ML	19
第三節 第二階段：驗證 SIS-ML 之心理計量特性	24
第六章 研究結果	26
第一節 第一階段：SIS-ML 之發展	26
第二節 第二階段：驗證 SIS-ML 之心理計量特性	27

第七章	討論	28
第八章	總結	32
參考資料	33
附錄一：SIS-ML 各向度之候選題目（依預測力順序）	62
附錄二：17 組 SIS-ML 候選題組之題目	63



圖目錄



圖一：決策樹之基本結構	39
圖二：腦神經細胞	40
圖三：人工神經元	41
圖四：人工神經網路（以 2 層隱藏層為例）	42
圖五：SIS-ML 之分數轉換流程	43
圖六：SIS-ML 發展流程	44
圖七：步驟 1 所使用之包含套索迴歸之 ANN 模型圖	45
圖八：SIS 各題目對各向度分數之預測程度（模型參數絕對值加總數值越大，預測力越高）	46
圖九：各 ANN 模型架構於各題組之預測力	47
圖十：27 題題組之 6X196 模型於測試資料 (n=77) 之模擬分數及原始分數之散佈圖	48
圖十一：SIS-ML 之 ANN 模型模擬分數示意圖（以單一個案分數為例）	49

表目錄



表一：SIS 用於中風患者之信度及反應性彙整	50
表二：SSQOL 用於中風患者之信度及反應性彙整	51
表三：以 NIHSS 及 BI 驗證 SIS 各向度收斂效度之數據彙整.....	52
表四：6X8 及 6X32 訓練資料 (n=89) 模型分數	53
表五：6X196 及 6X512 訓練資料 (n=89) 模型分數	54
表六：10X8 及 10X32 訓練資料 (n=89) 模型分數	55
表七：10X196 及 10X512 訓練資料 (n=89) 模型分數	56
表八：6X196、10X196 及 6X512 驗證資料 (n=90) 之模型分數.....	57
表九：6X196、10X196 及 6X512 之訓練 (n=89)、驗證 (n=90)、測試 (n=77) 資 料之模型分數	58
表十：27 至 32 題題組之訓練 (n=89)、驗證 (n=90)、測試 (n=77) 資料之 8 個向 度平均模型分數	59
表十一：SIS-ML 於測試資料 (n=77) 之平均值、標準誤及隨機測量誤差	60
表十二：SIS-ML 於測試資料之同時效度及收斂效度	61

第一章 中風患者之生活品質簡介

第一節 中風之病理學簡介

中風，又稱腦血管疾病，是造成全球人口死亡及失能之主要原因。在臺灣地區，依據衛生福利部民國 107 年國人十大死因統計顯示，腦血管疾病位居第四（行政院衛生福利部，2018）。中風患者常伴隨身體功能之損傷，包含生理、心理、認知功能，其中以動作損傷最為常見。身體功能損傷往往影響個案之日常生活執行能力，進而影響其健康生活品質。本章將簡介中風之流行病學、定義、分類與臨床症狀，並探討中風對個案之健康生活品質之影響，及臨床人員掌握個案健康生活品質之重要性。

中風之定義

根據 WHO 之定義，中風係指「源自於血管之突發性局部（或多處）神經性損傷，發病者症狀持續發作 24 小時或於 24 小時內身亡，且排除其他非血管原發性之病因」（World-Health-Organisation., 1978）。此定義適用於所有型態之中風類型。若個案之臨床症狀於 24 小時內緩解，則稱為「暫時性腦缺血發作 (Transient Ischaemic Attack)」，不包含於中風之定義內。

中風之流行病學

中風長期高居世界十大死因之一。據世界衛生組織 (World Health Organization, WHO) 2016 年之統計，每年全球約有 569 萬人因中風死亡，位居十大死因第二。其發生率於 55 歲以上之人口平均每千人約 4.2–6.5 人 (Emsley & Hopkins, 2008)。而在臺灣地區，36 歲以上人口之發生率約為 3/1000。盛行率方面，全球之盛行率約為 25/1000，臺灣在 36 歲以上之中風盛行率約為 16.42/1000 (Chang & Tseng, 2003; Hu et al., 1989; Hu, Sheng, Chu, Lan, & Chiang, 1992)。由於中風患者常伴隨不同程度之功能損傷，疾病影響病人本身所帶來之社會成本相當可觀。



中風病因之分類

中風可依據發生之原因分類：阻塞型與出血型。阻塞型中風 (Ischemic stroke) 為腦內或頸部血管阻塞，使腦細胞無法經由血液正常供氧持續三分鐘，導致該腦區缺氧壞死。阻塞原因可由血栓所致，亦可為身體內之異常雜質引起。當腦血管發生粥狀硬化（脂肪因堆積致使血管管壁內膜變厚及變硬，使血管失去彈性，導致血流量減少或出現血塊）時，血液中的血小板及纖維蛋白會聚集成血塊，以修補血管避免其失血或血流衝擊造成進一步傷害。但若血塊脫離，跟著血流移動，則形成血栓。此一血栓於動脈中流動，致使血管管腔日益狹窄，進而使血管發生阻塞。栓塞物亦可能是其他器官之碎片、組織片、脂塊、大團空氣栓或癌細胞等身體內之異常雜質。血管壁發生病變、血流速度減緩、血液黏稠度增加或紅血球過多時，都容易引起血管阻塞。出血性中風 (Hemorrhagic stroke) 為腦內血管發生出血現象，血塊壓迫腦組織，進而促使腦部組織壞死。依據出血位置不同而分成顱內出血 (Intracranial hemorrhage) 及蜘蛛膜下腔出血 (Subarachnoid hemorrhage)。顱內出血為大腦內腦血管破裂所致，若個案本身已有血管硬化，在忙碌、緊張、憤怒或出力時，容易引起血壓突然升高，造成出血。蜘蛛膜下腔出血則是介於蜘蛛膜及軟腦膜間空腔之血管破裂，常因腦動脈瘤破裂所致，是所有中風中致死率最高者 (Truelsen, Begg, & Mathers, 2006)。

中風之臨床症狀

中風的症狀依據腦組織受損部位不同，會表現出不同之神經學症狀。一般常見的中風症狀為突發性臉部、上肢或下肢的無力或麻痺，通常發生在身體單側。其他症狀包括：語言表達或理解困難、視力缺陷、行動困難、暈眩、平衡功能缺失、肢體協調障礙、不明原因之劇烈頭痛、暈厥或失去意識等 (Mackay & Mensah, 2004)。

中風患者具有極高之異質性。人體大腦結構中，不同區塊各自支配不同功能，因此不同病人會因損傷部位不同，表現出不同的症狀及嚴重性 (Gazzaniga, 2014)。不同區域之中風，會使得各腦區所掌管之功能受損，進而表現出不同的損傷症狀 (Emsley & Hopkins, 2008)。而發病範圍大小，則會影響中風病患有不同之損傷程度，發病範圍越大，功能損傷程度越嚴重，影響之功能層面也可能更廣

泛。部分病患可於中風後一段時間自行痊癒，但多數病患會造成永久性之功能損傷，留下不同嚴重程度之中風後遺症 (Pinter & Brainin, 2012)。



小結

近年來，實證醫學研究之發現及新型藥劑之治療，使全世界中風之死亡率及發生率逐年降低。然而，隨之而來之後遺症，成為中風患者必須面對之課題。中風造成失能之後遺症，是我國成人殘障的第一要因，更是使用健保資源前三名之疾病。以臺灣為例，每年約有一萬七千人會因中風導致日常生活失能，其中後續頻繁就醫也造成照顧者及家庭經濟沉重之負擔，嚴重影響中風患者之生活品質。

第二節 健康生活品質簡介

「生活品質 (quality of life, QOL)」之定義為從個人正負向之情緒、主觀認知層面以及身心健康的角度來評估個體整體之生活情形 (台灣版世界衛生組織生活品質問卷發展小組, 2000)。然而，不同領域的學者對於 QOL 有不同的看法及評量方式，如：經濟學家著重於國民所得，心理學家著重於個體之心理狀態及認知功能，醫師著重於生理功能及治療效益等。因此 QOL 又分為一般生活品質 (global QOL, QOL) 及健康相關生活品質 (health-relate QOL, HRQOL)。QOL 強調個人在所處環境中，對一般廣泛性的生活各方面之滿意度，常以個人之主觀感受作為評斷標準；HRQOL 則強調因疾病、意外事件或醫療處置所導致之個人身體功能改變，進而影響個體在生理、心理、社會、日常功能、安適感 (well-being) 層面之 HRQOL 的改變，可由主觀判斷及客觀測量來評量。近年來的醫療照護所提倡「以病患為中心 (patient-centered)」之照護不再只關注於專業人員之評估，病患個人主觀的自覺健康狀態，自覺生活滿意度等，都是重要的評估項目 (姚開屏, 2002)。

HRQOL 為病患主觀對各種健康層面所產生之主觀感受，有助於研究或臨床人員直接從病患之實際感受了解個案之健康狀態。依據評估對象可分成兩種：(1) 一般性 (generic) 及 (2) 特定疾病性 (disease-specific)。一般性評估為評估共通之生活品質，其評估結果可用來比較不同族群間的差異，但無法全面評估到特定族

群所特別關注之生活品質；特定疾病性評估則能深入瞭解特定族群之生活品質，但難以進行跨族群間之比較 (姚開屏, 2002)。

病患之 HRQOL 受生理、心理、認知、環境等多個健康層面影響。若能掌握個案各個 HRQOL 之向度，有助於掌握病患之自覺健康狀態，提供數據以供醫學療效分析、臨床決策分析、衛生政策之發展及個案之健康風險評估管理，以落實「以病患為中心」之概念。

第三節 評估中風患者之 HRQOL 之重要性

中風會導致患者在多個層面（如：生理、心理、社會及認知功能）受損。傳統臨床試驗多以中風患者的失能 (disability) 嚴重程度（如：修正版 Rankin 量表 (Modified Rankin Scale, MRS)、日常生活自理能力（如：巴氏量表 (Barthel Index, BI)）評估病人之預後 (Duncan et al., 1999; Williams, Weinberger, Harris, Clark, & Biller, 1999)。然而，以現今強調「以病患為中心」的整體照護及 ICF 強調的多向度評估，傳統之評估方式著重於身體功能的缺失和恢復的評量，缺乏對於其他向度（如：心理、社會、認知功能）之評估。

中風患者具有極高之異質性 (Emsley & Hopkins, 2008)。病患功能缺失常因發病位置、發病範圍不同，而可能造成不同部位之生理損傷。因此，不同個體之疾病嚴重程度、恢復進程、面對疾病之態度皆不相同。此外，中風患者常伴隨認知或語言功能缺失，導致與他人之溝通困難。上述中風患者之症狀影響個體之健康狀態。

HRQOL 以整體評估及主觀感受代表個案健康狀態，中風病患特定之 HRQOL 亦完整呈現中風患者在罹病後的全面健康狀態 (Guyatt, Feeny, & Patrick, 1993)。因此，掌握中風病人之 HRQOL，有助於釐清患者於治療進程之各階段之主觀感受，並且可協助臨床人員即時掌握患者於日常生活中全面性之健康狀況。

第二章 健康生活品質量表之介紹及評析

第一節 良好健康生活品質量表之主要特性

具備良好健康生活品質量表之主要特性有三：(1) 完整、明確之評估範疇；(2) 良好之心理計量特性 (psychometric property)；(3) 評估方式便利及快速。

一、完整、明確健康生活品質之評估範疇


完整、明確之健康生活品質評估範疇 (domains) 應含括個體之生理、心理與社會三大範疇。然而，各個範疇所包括之向度廣大，且不同疾病族群，著重之向度亦不相同。研究者統整出於中風患者中，HRQOL 常又細分 8 個向度：肌力、活動力、情緒、日常生活自理能力、社會參與能力、視力、溝通能力、認知功能 (Guyatt et al., 1993; Yin, Njai, Barker, Siegel, & Liao, 2016)。

二、良好之心理計量特性

心理計量特性為評判評估工具之評估結果是否為穩定、可信之指標 (Mokkink, Prinsen, Bouter, Vet, & Terwee, 2016)。評估工具須完整地驗證其心理計量特性，方能提供完整之數據，供臨床人員及研究人員參考。臨床人員使用心理計量特性良好之評估工具，方能掌握個案正確之能力、介入後之能力變化及預測患者之預後。研究人員亦須使用良好心理計量特性之評估工具，才能有效評估欲探討之特質所涵蓋之概念，進而發展有效之介入方案或作為發展其它評估工具之效標 (Kirshner & Guyatt, 1985; Mokkink et al., 2016)。因此，評估工具若有良好之心理計量特性，將可成為臨床人員及研究人員提升醫療品質及研究品質之利器。常見之心理計量特性有三：信度 (reliability)、效度 (validity) 及反應性 (responsiveness)(Mokkink et al., 2010)。

信度

信度為重複測量某一穩定特質時，獲得相同結果之程度 (Kline, 1998)。信度驗證包含再測信度 (test-retest reliability)、隨機測量誤差 (random measurement error) 及練習效應 (practice effect)。

- 
1. 再測信度：欲評估之特質為穩定特質之情況下，同一位患者於相同情境但不同時間接受同一測驗多次評估，所得分數一致之程度，即代表評量工具重複評量時分數穩定之程度 (Kline, 1998)。
 2. 隨機測量誤差：評估工具每次施測之結果都可能受到一些隨機產生、無法避免之原因影響，造成評估結果在一定範圍內波動。評估工具之隨機測量誤差愈小，其評估結果之穩定性及精準度愈高 (Atkinson & Nevill, 1998)。
 3. 練習效應：同一評估工具重複施測於同一個案時，個案測驗結果卻隨著次數增加而越來越好，但個案之實際能力未必有真實之改變，為系統性誤差。練習效應越小，代表評估結果之穩定性及精準度愈高。練習效應可能受欲評估之特質之穩定性影響（如：認知測驗普遍有練習效應）(Dodrill & Troupin, 1975)。

效度

效度指評量工具的評量結果可評估到該評估工具欲測量之特質（受測者之特定能力）或預測個體能力/表現之程度 (Kline, 1998)。重要之效度驗證概念包含內容效度 (content validity)、同時效度 (concurrent validity)、收斂效度 (convergent validity)、及生態效度 (ecological validity)。

1. 內容效度：指測驗內容是否涵蓋使用者欲測量之特質。因檢驗方法為聘請相關專家檢驗量表之內容效度，因此又稱專家效度 (Kline, 1998)。
2. 同時效度：指測驗結果與欲測量特質之公認測量方法（黃金標準）所得結果的關連程度，有助於推估評量工具所評量之特質是否與欲評量之特質相同。發展短版測驗可使用原版測驗當作同時效度之效標測驗 (Hobart, Lamping, & Thompson, 1996)。
3. 收斂效度：當欲測量之特質缺乏黃金標準時，可驗證測驗與理論上相關特質之其他評估工具之關連程度，有助於推估評量工具所評量之特質是否與欲評量之特質相同 (Portney, 2020)。
4. 生態效度：指測驗結果可類化個案於真實生活環境中之行為或功能表現之程度，即評量結果是否可作為推估患者真實生活能力高低依據的重要

指標 (Sbordone & Long, 1996)。

反應性

反應性表示評量工具可偵測個案於一段特定時間內功能變化之能力，即評量工具能否反應治療成效的重要指標 (Hobart et al., 1996)。

三、評估方式便利及快速

臨床環境中，分配於臨床評估之人力及時間有限。且臺灣之臨床環境中，中風個案之介入時間以 30 分鐘為單位，若評估工具之施測時間可縮短至 30 分鐘內完成，可行性將隨之增加。

第二節 常用於中風患者之 HRQOL 量表

HRQOL 量表為多向度之評估，測量分數有兩種解釋方式：(1) 整體之 HRQOL (overall HRQOL, O-HRQOL)：綜合評估病患整體之健康狀態；(2) 各向度之 HRQOL：剖析 HRQOL 各向度之健康狀態。整體 HRQOL 僅呈現單一數值，較為單純，易於使用者解讀分數。各向度之 HRQOL 則能反應出所有向度之健康狀態 (Williams et al., 1999)。

常用於中風患者專屬之 HRQOL 量表有二：(1) 中風衝擊量表 3.0 (Stroke Impact Scale 3.0, SIS)；(2) 中風專屬生活品質量表 (Stroke-Specific Quality of Life, SSQOL)。

中風衝擊量表

SIS 為專屬評估中風患者之 HRQOL 量表 (Duncan et al., 1999)。SIS 包含 8 個向度：肌力 (strength)、活動力 (mobility)、情緒 (emotion)、日常生活自理能力 (ADL/IADL)、手部功能 (hand function)、溝通能力 (communication)、記憶力 (memory) 及社會參與能力 (participation)，初版為 64 題。目前版本為分析題目難度後，將部分題目刪除後得到 SIS 3.0 的版本，共 59 題。SIS 計分方式可分成各向度分數加總及量表總分加總量表之計分方式，每個向度總分皆為 100 分，單一向度分數越高代表單一向度 HRQOL 越佳，整體分數越高代表 O-HRQOL 越佳。SIS 已在多個國家（如：美國、日本、韓國、德國、澳洲）之中風患者驗證心理計量特性 (Carod-Artal, Coral, Trizotto, & Moreira, 2008; Choi et al., 2017; Geyh,



Cieza, & Stucki, 2009; Kamwesiga, von Koch, Kottorp, & Guidetti, 2016; Mohammad, Al-Sadat, Siew Yim, & Chinna, 2014; Ochi, Ohashi, Hachisuka, & Saeki, 2017; Vellone et al., 2015)。臺灣亦有中文翻譯版，也已驗證於中文版於臺灣中風患者之心理計量特性 (Chou, Ou, & Chiang, 2015; Lin et al., 2010; 梁佩蓉 & 林佩欣, 2016)。信度方面，除情緒部分有中度之信度以外，其他 7 個向度具有高度之信度。效度部分，日常生活自理能力、活動力、記憶力與溝通能力皆有良好之收斂效度，肌力、手部功能及情緒有中度之收斂效度，社會參與能力則有低度至中度之收斂效度。反應性部份，SIS 有效果值中等至大之反應性 (Chou et al., 2015; Lin et al., 2010; 梁佩蓉 & 林佩欣, 2016)。

中風專屬生活品質量表

SSQOL 為專為中風病患設計且以病患為中心發展之 HRQOL 問卷 (Williams et al., 1999)。原版 SSQOL 包含 12 個 HRQOL 之向度：行動 (mobility)、精神體力 (energy)、語言 (language)、心情 (mood)、個性 (personality)、自我照顧 (self-care)、社會角色 (social roles)、家庭角色 (family roles)、思考 (thinking)、上肢功能 (upper extremity function)、視力 (vision) 及工作 (work)，共 49 題。另有刪題後之短版，合併自我照顧、行動、上肢功能及工作/生產力為 1 個向度，共包含 8 個向度，共 47 題。評分方式為 1-5 分，依評估項目個案回答之選項可分為 3 種：(一)「完全需要幫忙」至「完全不需要幫忙」；(二)「完全無法做到」至「完全沒有困難」；(三)「非常同意」至「非常不同意」，1 分對應患者完全需要幫忙/完全無法做到/非常同意，5 分對應患者完全不需要幫忙/完全沒有困難/非常不同意。SSQOL 計分方式可分成各向度分數加總及量表總分加總，總分範圍為 12-60 分。單一向度分數越高代表單一向度 HRQOL 越佳，總分越高代表 O-HRQOL 越佳。SSQOL 亦於多個國家 (如：德國、丹麥、土耳其、波斯語系國家等) 之中風患者驗證心理計量特性 (Chen, Wu, Lin, Li, & Yu, 2012; Cruz-Cruz et al., 2013; Ganvir, Harishch, & Kunde, 2018; Hakverdioglu Yont & Khorshid, 2012; Legris et al., 2018; Mahmoodi et al., 2015; Muus, Williams, & Ringsberg, 2007; Odetunde, Akinpelu, & Odole, 2017, 2018; Pedersen et al., 2018; Sallam, Al-Khamis, Muaidi, & Abdulla, 2019)。臺灣亦有中文翻譯版，也已驗證於中文版於臺灣中風患者之心理

計量特性 (Chou et al., 2015; Hsueh, Jeng, Lee, Sheu, & Hsieh, 2011; Lin et al., 2010)。信度方面，SSQOL 有良好之信度。效度部分，SSQOL 僅有低度至中度之收斂效度。反應性方面，SSQOL 僅有效應值小之反應性 (Chou et al., 2015; Hsueh et al., 2011; Lin et al., 2010)。

第三節 比較 HRQOL 量表及發展短版量表之選擇

SIS 與 SSQOL 皆是中風患者專屬之 HRQOL 量表，且都是涵蓋多向度之量表，因此昔日研究曾比較兩者之心理計量特性。研究結果整理於表一及表二。

SIS 之各向度有良好之內在一致性 ($\alpha = 0.71-0.98$)。再測信度達中度至良好 ($ICC = 0.48-0.97$; $\rho = 0.50-0.88$)。隨機測量部分，有可接受至大之隨機測量誤差 ($SRD\% = 12\%-81\%$) (Chou et al., 2015)。反應性部分，有低度至中度之反應性 ($SRM = 0.04-0.52$, Wilcoxon $Z = 0.61-4.24$) (Lin et al., 2010)。

SSQOL 內在一致性達中度至良好 ($\alpha = 0.65-0.97$)。再測信度部分，各向度達中度至良好 ($ICC = 0.70-0.97$; $r = 0.36-0.96$; $\rho = 0.35-0.94$)。隨機測量部分，SSQOL 達可接受至大之隨機測量誤差 ($SRD\% = 19\% \sim 84\%$) (Chou et al., 2015)。反應性部分，各向度之反應性皆不佳 ($SRM = 0.01-0.17$, Wilcoxon $Z = 0.01-1.87$) (Lin et al., 2010)。

簡言之，SIS 具有較佳之再測信度、隨機測量誤差及反應性，較 SSQOL 更適合應用於研究或臨床環境。之前已有學者發展短版 SIS (SIS-16, SF-SIS) (Duncan, Lai, Bode, Perera, & DeRosa, 2003; MacIsaac et al., 2016)，然而，這些短版測驗皆採單一分數計分，僅能呈現整體 HRQOL，無法反應出各向度之 HRQOL 分數。因此，若能將機器學習應用於 SIS 3.0 之簡化，則有潛力發展短版之多向度 SIS，以利提升研究或臨床環境評估之效能。

第三章 運用機器學習模型發展短版量表



第一節 機器學習簡介

機器學習 (Machine Learning) 為人工智慧 (artificial intelligence, AI) 之子領域。AI 為訓練電腦或機器具有學習能力，使其做出之行為模式或判斷近似人腦之結果，如：決策、問題解決、物品辨識及特徵萃取。其原理為訂立好規則或數學模型後，讓電腦能夠根據所輸入之資料輸出判斷的結果。機器學習藉由大量數據中之資訊及知識，提高模型預測之性能，並做出適當之決策。因機器學習涉及數據統計，因此，機器學習是基於統計學，優化數學理論、矩陣代數及微積分，用以實現電腦科學及工程學預測概念之理論。與統計模型不同之處，機器學習不需要對資料進行假設，且機器學習之目的並非描述及驗證數據與模型輸出變數之關係。機器學習之目的在於「訓練」模型，使模型獲得最佳性能，藉以獲得可反覆預測之模型。依據學習方式不同，機器學習可分為監督式學習、非監督式學習及強化學習 (LeCun, Bengio, & Hinton, 2015; Talabis, McPherson, Miyamoto, & Martin, 2014)。

一、監督式學習

使用監督式學習，資料需包含大量原始數據及標準答案。監督式學習常處理兩種問題：迴歸及分類。若預測目標為連續資料時，則處理迴歸問題；若目標為離散資料時，則處理分類問題。在訓練模型時，技術人員需先將資料分為訓練資料 (training set) 及測試資料 (testing set)。其次，再使用訓練資料細分出驗證資料 (validation Set)，以利進行交叉驗證 (cross validation)，確保模型是否出現過度擬合 (over fitting) 或選擇偏差 (selection bias)。訓練模型之步驟，首先需使用訓練資料透過特定之演算法選取使用於模型之資料屬性，建構出適當之機器學習模型。其次，再利用驗證資料比對已訓練好之機器學習模型其演算結果與答案之一致性。若計算結果與答案之一致性不佳，則再調整模型之所有參數，並重複上述之步驟。模型訓練完成後，再使用訓練資料檢驗最後模型之結果，判斷模型之效能。以預測個案預後之模型為例，該模型使用已知數天後個案之功能性評估之資

料庫做為監督式學習結果之答案，利用監督式學習訓練模型，並驗證模型預測個案功能性評估結果與答案之一致性，以預測個案之預後。

常用於處理連續變項之監督式機器學習模型有二：(1) 支援向量機 (Support Vector Machine, SVM)、(2) 決策樹 (Decision Tree)。

支援向量機 (Support Vector Machine, SVM)

SVM 原是用於處理離散資料之分類模型，用以處理屬性分類 (classification) 之問題。其原理為運用統計風險最小化之原則來估計一個分類的超平面 (hyperplane)。SVM 模型可將自變項及依變項之對應關係自低維度向量空間提升至高維度之向量空間，使資料於不同類別之分類達最佳化。隨著人工智慧技術之提升，SVM 目前可應用於連續資料，用以找出準確預測資料之分布平面。

決策樹 (decision tree)

決策樹來自於決策理論，由一個決策圖和可能的結果（包括資源成本和風險）組成，目的為輔助決策，常用於決策分析時用以創建目標之規劃。決策樹的理論基礎在於計算條件概率以建構預測模型，他代表的是對象屬性與對象值之間的一種映射關係。因其決策模型類似樹狀之決策支持工具（包含隨機事件結果，資源代價和實用性），因此稱作決策樹。依據處理數據類型不同，可分為分類樹及迴歸樹。構成決策樹之元素包含內部節點 (Internal Node)、分支 (Branch)、樹葉節點 (Leaf Node) 及根節點 (Root Node)。決策樹自根節點為起始點，樹中每個內部節點代表對某個對象（或屬性）之測試，而每個節點之間的分支代表其對應測試之結果，每個樹葉節點則對應從根節點到該樹葉節點所經歷的路徑所呈現之結果（圖一）。決策樹僅有單一輸出，若欲有複數輸出，可以建立獨立的決策樹以處理不同輸出。

二、非監督式學習

非監督式學習目的為處理未標記之數據或未知結構之數據，以探索資料之結構，屬探索性數據分析技術。非監督式學習不需準備標準答案，最常使用之技術為聚類分析 (Clustering)：使用未分類或標記之大量數據中，尋找其相關性進行分組，常應用於市場區隔之分析，如：透過聚類分析市場中顧客之行為，以便制定不同聚類的顧客不同的行銷策略。



三、強化學習

強化學習目的為使模型可透過每次對於環境之互動，動態調整模型參數，使模型得以取得最大化之利益，以完成特定任務。強化學習與監督式學習不同之處在於強化學習不需要準備正確答案，因強化學習更著重於動態規劃，需要於探索未知領域及遵從現有知識中獲取平衡，以增強學習效能，發揮其最大性能。國際象棋即是使用強化學習訓練之例子，強化學習即是透過即時與延遲反饋完成模型之學習。

第二節 人工神經網路

人工神經網路 (artificial neural network, ANN) 為機器學習其中一種方法，簡稱神經網路或類神經網路。ANN 是一種模仿生物神經網路結構之計算模型，用以對函式進行估算或近似。ANN 透過模擬腦神經細胞之運作模式，模仿人的大腦來處理輸入信號並將其轉換為輸出信號。它提供功能強大的建模算法，可實現特徵變量與輸出信號之間的非線性關係。ANN 可從數據中學習，無需事先對於特定功能做假設，適用於當研究人員不了解其基本功能的複雜現象時使用。ANN 亦包含監督式學習、非監督式學習及強化學習三種訓練模型之方法。

一、腦神經細胞

構成腦神經細胞之元件為：神經細胞核 (nucleus)、軸突 (Axon)、樹突 (Dendrites)、突觸 (Synapse) (圖二)。

1. 神經細胞核 (soma)

是神經元蛋白質合成的主要場所。樹突將收集到的信號傳遞至此後做一次非線性轉換，再由軸突將神經信號傳送至其他神經細胞。

2. 軸突 (Axon)

連接於神經細胞體之突出，功能為傳遞細胞本體之動作電位至突觸，為主要神經信號傳遞渠道。

3. 樹突 (Dendrites)

為神經元的輸入通道，其功能是将自其他神經元所接收的動作電位 (電信



號) 傳送至細胞本體。其他神經元的動作電位藉由位於樹突分支上的多個突觸傳送至樹突上。

4. 突觸 (Synapse)

是神經元之間，或神經元與肌細胞、腺體之間通信的特異性接頭。中樞神經系統中的神經元以突觸的形式互聯，形成神經元網絡。突觸亦是中樞神經系統和身體的其它部分（如：肌肉和各種感受器）交換信息的渠道。

二、人工神經元

人工神經元之概念來自於生物之神經元。構成人工神經元之元件包含：輸入值、加權值 (weight)、向量總和、激活函數 (active function)、輸出值。人工神經元之結構示意圖如圖三所示。

1. 輸入值 (x_1, x_2, \dots, x_n)：對應樹突傳遞之神經訊息，每一個人工神經元皆有多個輸入值，為純量。
2. 加權值 (w_1, w_2, \dots, w_n)：對應每個樹突與神經細胞核之連結強度，每一個輸入值皆對應一個加權值，為純量。
3. 權值總和 (Σ)：對應傳遞至神經細胞核之神經訊號，為加權後的輸入值總和，為純量。
4. 激活函數 ($\varphi(\theta)$)：用以對權值總和做一非線性轉換，用以模擬神經系統之閾值，以決定訊號是否往後做傳輸。傳統之激活函數包含 Sigmoid 函數、雙曲正切 (tanh) 函數 (hyperbolic tangent function)。深度學習網路中常會使用 ReLU 函數 (Rectified linear unit)，以達到更好之學習效果。
5. 輸出值 (y_k)：對應傳遞至軸突之神經訊號，為激活函數轉換後之純量。轉換後之數值介於-1 到 1 之間。

每一個人工神經元皆有多個輸入值 x_1, x_2, \dots, x_n 及輸出值 y_k ，輸入值與輸出值的關係式，一般可用輸入值的加權乘積和的函數來表示。

三、人工神經網路

人工神經網路是由數個人工神經元經由縱向或橫向擴展之網路。人工神經網路結構分為接收訊息之輸入層 (input layer)、輸出答案之輸出層 (output layer)，輸入層與輸出層之間「隱藏」了數層無法看見之作用機制，因此稱為隱藏層

(hidden layer)。每層（除了輸入層以外）之輸入值為上一層神經元之輸出值，經過權值加總後，再經過該層人工神經元之激活函數，傳出輸出值至下一層。ANN 必有一層輸入層、一層輸出層及一層以上之隱藏層。人工神經網路之結構示意圖如圖四所示。

「目標函數」為優化機器學習模型之算則。當資料與模型進行擬合時，還需要先使用目標函數以進行模型之優化。目標函數有兩個非常重要的元素：(1) 損失函數 (Loss function)：衡量預測值與實際值的差異，讓模型能往正確的方向學習；(2) 正則化 (Regularization)：避免模型變得過於複雜，造成過度擬合 (overfitting)。迴歸模型常會使用均方誤差 (mean-square error, MSE)、平均絕對誤差 (mean absolute error, MAE) 作為損失函數。為了避免過度擬合，我們可以把正則化加入目標函數中，以懲罰 (penalty) 模型的複雜度，當模型愈複雜，其值就愈大。正則化的損失函數分為兩種：

1. L1 正則化損失函數：指權值矩陣 w 中各元素之絕對值之和，亦為正則化 MAE 損失函數。L1 正則化可以產生稀疏權值矩陣（矩陣 w 會變得較稀疏，意思是 w 向量有很多的 0 在裡面），即產生一個稀疏模型，可以用於特徵選擇 (feature selection)，以利簡化模型。L1 亦可以防止模型過度擬合，優點為較省記憶體，缺點為防止過度擬合效果不及 L2 好。實現 L1 損失函數的線性迴歸模型稱為套索迴歸 (lasso regression)。

2. L2 正則化損失函數：指權值矩陣 w 中各元素之平方和後再求平方根，亦為正則化 MSE 的損失函數。L2 正則化可防止模型過度擬合。實現 L2 損失函數的線性迴歸模型稱為嶺迴歸 (ridge regression)。

第三節 機器學習之實作

機器學習為運用演算法使機器「學習」之電腦科學。其目的為讓電腦能夠自行提升學習能力，可藉由增加已知資料而提高預測未知資料之準確性，以節省使用者人工調整校正之精力。典型的應用包含概念學習 (concept learning)、函數學習 (function learning)、預測模型 (predictive modeling)、聚類 (clustering) 與找尋預測特徵 (finding predictive patterns)。

Scikit-learn (Sklearn) 為 Python 實作機器學習的套件之一，是通用機器學習開源資料庫 (Pedregosa et al., 2011)。Sklearn 之優點在於使用者亦於入門，對於此開源軟體下之任何一種分類或迴歸模型皆可使用相同之方法處理，無須在應用程式介面 (Application Programming Interface, API) 進行鑽研。SKlearn 提供應用機器學習之六大功能：資料的前置處理 (preprocess)、監督式學習的分類 (classification) 以及迴歸 (regression) 演算法、非監督式的聚類演算法 (clustering)、維度約減 (dimensionality reduction) 及模型選擇 (model selection)。完整的機器學習實作流程為：(1) 進行資料轉換：包含資料之前置處理、結構分析及特徵提取，掌握資料之訊息，使資料進行清理及變換，以利用於機器學習之演算法；(2) 選擇適當之學習演算法：包含監督式學習之分類演算法、迴歸演算法及非監督式學習之聚類演算法；(3) 模型選擇：通過調節參數以提升模型準確率，經比較及驗證後，判斷模型之好壞。

Keras 為一 Python 所編寫的開源神經網路庫，可於多個平台上執行。Keras 旨在快速實現 ANN 之網路架構，較 Sklearn 有更進階之功能，使用者可自行設計及調整 ANN 之任何參數 (各隱藏層所使用之激活函數及損失函數、正則化及優化等)。本研究需使用第一層 ANN 模型包含 L1 正則化損失函數，因此使用 Keras 之 ANN 模型架構以進行特徵題目之篩選。

多層感知器 (Multi-layer Perceptron, MLP) 為 Sklearn 及 Keras 套件中實現 ANN 之網路架構，因此，MLP 網路架構包含輸入層、隱藏層及輸出層。進行 MLP 模型之訓練需進行兩種參數之決定：(1) 目標函數之種類及參數；(2) 隱藏層之層數及神經元個數。目標函數的決定取決於研究問題，亦可藉由比較預計使用之 ANN 大小、訓練時間、模型精確度及硬體需求來決定。輸入層及輸出層之

神經元已於擬合資料時決定，輸入層為預測因子（或控制及影響因子），輸出層為輸出結果。因此，使用者僅可設定隱藏層之層數及各隱藏層之神經元數。隱藏層之神經元數量可決定 ANN 參數多寡及影響 ANN 歸納推演之能力，因此，神經元個數過少無法建構適當之模型以描述問題；相反地，過多的神經元個數因自由度過高，以至於模型過度擬合特定資料，在使用其他資料擬合模型時反而容易失去模型應有之歸納推演能力。

第四節 運用「機器學習模型」發展短版測驗之優勢

機器學習的主要目的之一在於訓練模型獲得最佳預測性能 (Senders et al., 2018)。研究發現機器學習可不限於線性關係之預測。因此，機器學習可由大量數據中之資訊及知識中提取重要之特徵，習得可精準預測預後的模型參數，有效提升傳統線性迴歸模型的預測精準度，並做出適當之決策 (Senders et al., 2018)。相較於傳統之統計模型，機器學習所使用之資料不限於特定分布，且機器學習更著重於獲得可預測之模型，因此可自行由大數據資料庫中發現變項之間複雜之關係。

發展短版測驗之目的在於提升原版測驗之評估效能，以利應用於研究或臨床環境。短版測驗篩選原版之題目，提取欲評估之特質最相關之題目，降低評估時間及評估人力。過去研究多使用統計模型以發展短版測驗。然而，統計模型仰賴研究人員對於各個變項之掌握程度，且統計模型受限於資料須符合特定之分布，才能建構精準之統計模型。過去研究曾發展中風衝擊量表短版測驗 (SIS-16, SF-SIS) (Duncan et al., 2003; MacIsaac et al., 2016)。然而，昔日研究之短版皆使用單一總分計分，且 SIS-16 於刪減題目時未考慮所刪減之題目是否包含原版之 8 個向度。因此，目前尚無包含 8 個向度之 SIS 短版測驗。

機器學習之優勢在於可從大數據資料中學習並提取特徵，使用不同之學習方式訓練模型以達成任務 (Guyon & Elisseeff, 2003)。目前已有影像研究將機器學習模型應用於辨識影像之診斷系統，亦有研究使用機器學習模型發展人格評估 (Bleidorn & Hopwood, 2019; Deo, 2015; Giger, 2018)。然而，昔日研究多使用機器學習模型以完成分類任務（類別分析），鮮少有研究使用機器學習之迴歸任務（迴歸分析）發展或改良評估工具。若可使用機器學習之特性，應用於發展評估

工具，機器學習模型將有潛力應用於提取評估該向度最重要之題目，可有效降低發展短版評估工具之人力及時間成本，且能保有原版量表之理論架構/向度。因此，運用機器學習模型之預測能力將有潛力發展短版之多向度測驗。



第四章 研究目的

為解決目前 SIS 之各項限制，本研究計畫包含二個階段。第一階段之目的為：運用機器學習模型，發展短版 SIS (Stroke Impact Scale-Brief Machine learning version, SIS-ML)。研究者預期機器學習模型可有效運用 SIS-ML 模擬原版 SIS 之分數，且 SIS-ML 可精準評量中風患者之生活品質程度。SIS-ML 之分數轉換流程如圖五所示。

第二階段之目的為：驗證 SIS-ML 之心理計量特性。研究者將驗證 SIS-ML 之各向度分數應用於中風患者之同時效度及收斂效度。研究者預期 SIS-ML 可與原版 SIS 有高度相關性，有良好之同時效度；SIS-ML 之各向度與 NIHSS 及 BI 可有適當之相關程度：於肌力、活動力、日常生活自理能力、手部功能及社會參與能力可達中度以上之相關性，於情緒、溝通能力、記憶力可達低度相關性。

第五章 研究方法



本研究方法分為二階段：(一)發展 SIS-ML 及 (二)驗證 SIS-ML 之心理計量特性。本研究為次級資料分析，藉由國內昔日研究之收案資料作為發展 SIS-ML 及驗證其心理計量特性之基礎。本研究之初步設計構想如圖六。

第一節 資料來源及研究對象

研究者使用國內昔日研究之資料作為模擬分析之樣本 (Chou et al., 2015)，該樣本於 2008 至 2010 年，收案自台灣五所醫院，共 263 人，且皆有簽屬同意書。該研究之納入條件為：(1) 依據 DSM-IV 之診斷標準，診斷為中風、(2) 因中風而有單側偏癱、(3) 年齡超過 18 歲、(4) 具有足夠之閱讀及聽覺理解能力以完成自評生活品質量表、(5) 具備可完成單步驟指令之認知功能。該資料之人口學資料包含個案之性別、年齡、教育、婚姻狀況、工作狀況、診斷及中風次數。使用之量表包含美國國家衛生研究院中風量表 (National Institutes of Health Stroke Scale, NIHSS)、巴氏量表 (Barthel Index, BI) 及中風衝擊量表 3.0 版 (Stroke Impact Scale 3.0, SIS)。研究者將依照 35%、35%、30% 之比例，將樣本分為訓練資料 (training set)、驗證資料 (validation set) 及測試資料 (testing set)，以利進行交叉驗證。訓練資料用以訓練模型，驗證資料用以驗證訓練之模型是否可類化至其他資料，測試資料用以驗證短版題目之模擬分數與原始分數之預測性，以比較不同模型並選擇最適合預測之模型。

第二節 第一階段：發展 SIS-ML

研究者將依據下列五個步驟發展 SIS-ML：(1) 挑選短版題目之題組；(2) 訓練 ANN 模型以預測各向度分數；(3) 挑選預測力較佳之 ANN 模型架構；(4) 挑選預測力較佳之題數組合；(5) 選擇最適之機器學習模型。



步驟 1：挑選短版題目之題組

題組總數及題目組成

研究者組成 16 題至 32 題共 17 組題數不等之後選題組。候選題組 16 題與 32 題之題組為自 8 向度中，每向度提取 2 題及 4 題之題數均等之題目。17 題至 31 題共 15 組題組為 8 個向度題數不均等之題組。

挑題規則

研究者使用之挑題規則有二：(1) 短版題目須包含 8 個向度之題目；(2) 各向度題目至少須包含 2 題，至多包含 4 題。

程序

研究者使用訓練資料及 Keras 演算法 (Van Merriënboer et al., 2015)，匯入原版 SIS 之 59 題作為輸入層，並使用包含 L1 損失函數 (凸顯各題之預測能力) 之 ANN 模型，計算每題對於 8 個向度的個別分數之預測能力並做排序。之後分別自各向度中手動挑選出預測力最高的前 4 題作為特徵題目，並將所有選出之題目分別組成 32 題之短版題組。之後再依據 32 題之預測能力做排序，取出各向度題目預測力排名第 3、4 之題目之 16 題做為選題之依據，依序排除預測力較低之題目，做成 17 題至 31 題之題組。共組成 17 組短版題目之題組。

資料分析

研究者使用 Keras 之 4 層隱藏層、32 個人工神經元之 ANN 模型進行題目預測度之分析，該模型之第一層隱藏層加入 L1 損失函數作為選題依據。L1 函數之 α 值設定為 0.05，以凸顯各題目預測能力之差異。模型參數越大代表該題於 SIS 中之預測性越強，越有潛力做為短版題目 (Fonti & Belitser, 2017; Tibshirani, 1996)。研究者匯出 ANN 模型之模型參數，並取 59 題輸入層對應於各層之 32 個人工神經元之權值，取各題之模型參數絕對值加總值作為該題目對於原始 SIS 分數之預測力並排序 (Alhamzawi & Ali, 2018; Tozlu et al., 2020)，之後取出各向度排序前 4 及前 2 之題目，組成 32 題及 16 題之題組。之後再依據權值絕對值加總之順序選出 17 題至 31 題題組之題目。



步驟 2：使用不同模型架構訓練以短版題目模擬原版分數之 ANN 模型

ANN 模型

研究者使用 8 種 ANN 模型架構，並匯入訓練資料以訓練 ANN 模型。ANN 模型架構包含 2 個參數：隱藏層數及人工神經元個數。之前研究顯示適合用於分類或迴歸之 ANN 模型之隱藏層數為 5~10 層 (Senders et al., 2018)，人工神經元則多使用 2 之對數 (Jain, Mao, & Mohiuddin, 1996; Zhang, 2016)。本研究所使用之 ANN 模型參數分別設定 6 層及 10 層隱藏層數及 8、32、196、512 個人工神經元個數，組成 6X8、6X32、6X196、6X512、10X8、10X32、10X196、10X512 共 8 種 ANN 模型架構。

程序

研究者使用訓練資料進行模型訓練及模擬分析。研究者以 17 組選用之題目分數做為模型之輸入值，以各向度之模擬總分作為輸出值，以模擬原版 SIS 之各向度分數。17 組題目分別有不同之 ANN 模型以獲得其模擬分數，最終預計獲得 136 個機器學習模型及模擬分數結果。

研究者使用訓練資料訓練 ANN 模型後，對照訓練資料之模擬分數與原始分數之模型分數。

資料分析

研究者將使用多層感知器迴歸演算法 (Multi-layer Perceptron Regression, MR) 進行模型訓練。欲調整之參數包含：隱藏層之層數、隱藏層之人工神經元個數、激活函數、模型最大訓練次數 (影響模型精準度)，其餘參數使用演算法之預設函數。隱藏層將設定為 6 層及 10 層隱藏層數，人工神經元個數設定為 8、32、196、512，激活函數為整流線性單位函數 (Rectified Linear Unit, ReLU，為 ANN 常用之激活函數)， α 值設定為 50，模型最大訓練次數為 50000 次。

模型優劣以模型分數作為判斷指標。模型分數為多層感知器迴歸演算法判斷模型擬合資料優劣程度之指標，以判斷模型預測資料之預測力。其計算方式以相關指數 R^2 (又稱決定係數) 作為量化之依據。 R^2 之意義為以 SIS 總分之平均值作為評判依據，判斷 ANN 模型模擬之分數與原始 SIS 分數之一致性，通常用以判



定迴歸模型之優劣。 R^2 之公式如下：

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2}$$

1. \hat{y}_i ：ANN 模型預測之 SIS 分數。
2. y_i ：原版 SIS 分數。
3. \bar{y} ：訓練資料之原版 SIS 平均分數。

步驟 3：使用訓練資料挑選預測力較佳之 ANN 模型架構

程序

研究者比較步驟二所訓練之 136 組 ANN 模型，挑出模型分數較佳之 ANN 模型架構。

資料分析

研究者將對照 ANN 模型之模擬分數與原始分數之 R^2 ，作為分析模型優劣之指標。 $R^2 > 0.8$ 代表 ANN 模型有較好之模擬原版 SIS 分數之能力。最後比較 136 個模型之模型分數，並挑出 8 個 ANN 模型架構中，於各題組之 R^2 皆大於 0.8 之 ANN 模型架構，以進行下一步驟之分析。

步驟 4：使用驗證資料挑選預測力較佳之題數組合

程序

研究者使用驗證資料驗證步驟三保留之模型是否適用於其它資料，並確認模型是否有過度擬合訓練資料之情況。驗證資料之模擬分數皆以 R^2 對照實際分數並判斷模型之優劣。研究者比較驗證資料之 R^2 ，確認各模型可類化至其他資料之程度，並保留較佳題組。

資料分析

研究者使用步驟三篩選好之 ANN 模型，匯入驗證資料之選定題目之分數。模型匯出之模擬分數將對照原始分數之 R^2 ，作為分析模型優劣之指標。 $R^2 > 0.8$ 代表 ANN 模型於驗證資料中仍有較好之模擬原版 SIS 分數之能力。最後比較驗證資料之模型分數，並保留 8 個向度之 R^2 皆大於 0.8 之題組，以進行下一步驟之分析。



步驟 5：使用測試資料選擇最適之機器學習模型

程序

研究者於步驟四確認模型後，將匯入測試資料，並確認模型是否於測試資料仍有良好之一致性。最後，研究者比較各模型於訓練資料、驗證資料及測試資料於 8 個向度 R^2 之平均值，以挑選整體預測力最佳之模型（層數、每層神經元個數與 SIS-ML 總題數）。

資料分析

研究者使用步驟四所選定之 ANN 模型，匯入測試資料之選定題目之分數。模型匯出之模擬分數將對照原始分數之 R^2 ，作為分析模型優劣之指標。 $R^2 > 0.8$ 代表 ANN 模型於測試資料中仍有較好之模擬原版 SIS 分數之能力。研究者將比較測試資料之 R^2 ，確認各模型可類化至其他資料之程度，並計算訓練資料、驗證資料及測試資料 R^2 之平均值，以比較出最佳之 ANN 模型及題組。研究者亦繪製選定之題組及 ANN 模型之測試資料模擬分數及原始分數之散佈圖，並將原版 SIS 與 SIS-ML 視為相依資料，計算原版 SIS 與 SIS-ML 之平均值、標準差及原版 SIS 與 SIS-ML 分數差異絕對值之平均值及標準差，差異百分比為原版 SIS 與 SIS-ML 之差異絕對值平均值與原版 SIS 之平均值之百分比，以確認模擬分數及原始分數之趨勢及模擬分數之誤差。差異百分比 $< 10\%$ 代表 SIS-ML 與原版 SIS 之誤差為可接受之範圍 (Freedman, Pisani, & Purves, 2007)。

評估工具

中風衝擊量表 3.0 版 (Stroke Impact Scale 3.0, SIS)

本研究使用張谷州醫師所翻譯之 SIS 3.0 中文版。SIS 包含 8 個向度：肌力、記憶力、情緒、溝通能力、日常生活自理能力、活動力、手部功能及社會參與能力，每個向度最少 4 題，最多 10 題，共計 59 題。SIS 採 Likert 五分法計分，以評估個體對生活中各面向之滿意程度及重要感受。各向度獨立計分，除情緒向度有 3 題反向題，其它向度得分越高，表示個案自覺生活品質越佳。

第三節 第二階段：驗證 SIS-ML 之心理計量特性

研究二之目的為驗證 SIS-ML 應用於中風患者之心理計量特性。研究者將驗證 SIS-ML 各向度分數應用於中風患者之同時效度及收斂效度。



參與者

以次級資料分析之方法，研究者使用研究一所使用之測試資料，分析 SIS-ML 之同時效度及收斂效度。研究者將保留完成所有人口學資料、原版 SIS、美國國家衛生研究院中風量表 (National Institutes of Health Stroke Scale, NIHSS) 及巴氏量表 (Barthel Index, BI) 之受測者資料。

程序

本研究使用研究一所選定之短版題組及 ANN 模型，模擬測試資料之原版 SIS 分數，並且分析模擬分數與原版 SIS、NIHSS 及 BI 之相關性，以驗證 SIS-ML 之同時效度及收斂效度。符合收案標準之中風患者皆須完成原版 SIS、NIHSS 及 BI 之所有測驗。

評估工具

同時效度效標測驗

本研究使用研究一所使用之原版 SIS 作為同時效度之效標測驗。

收斂效度效標測驗

美國國家衛生研究院中風量表 (National Institutes of Health Stroke Scale, NIHSS)

NIHSS 系美國國家衛生研究院於 1980 年代所見，可對中風病患進行整體嚴重度之客觀評量 (Lyden et al., 1994)。NIHSS 內容涵蓋重要之神經功能評估，可將評估內容作系統性的量化。評估者需受過完整之訓練，方能進行評估。NIHSS 共計 11 個向度，包含：意識程度、眼球運動、視野、面部肌力、上肢運動、下肢運動、肢體運動失調、感覺功能、語言功能、構音困難及忽略。評估者依受測者之實際表現計分，總分為 0–42 分，分數越高代表中風嚴重度越高。個體之中風嚴重程度依照分數切截點分級：總分 3 分以下為輕微損傷，4–6 分以下為輕度

損傷，7–15 分為中度損傷，16 分以上為嚴重損傷。NIHSS 於國內、外研究結果顯示有良好之心理計量特性。

巴氏量表 (Barthel Index, BI)

BI 為國內外廣泛使用於評估基本日常生活功能 (Activities of Daily Living, ADL) 之指標，可評估個體對於日常生活功能之獨立程度 (Collin, Wade, Davies, & Horne, 1988)。BI 總計 10 題，內容包含進食、個人衛生、上廁所、洗澡、穿脫衣服、大便控制、小便控制、平地行走、上下樓梯及上下床或椅子，總分為 0–20 分。分數越高代表個體之基本日常生活功能越好。BI 於中風族群有良好之心理計量特性。

資料分析

同時效度

同時效度以皮爾森相關係數 (Pearson correlation coefficient, r) 分別分析 SIS-ML 各向度之總分與效標測驗分數之相關程度。同時效度之效標測驗為原版 SIS。 $r \geq 0.75$ 為良好； $0.40-0.74$ 為中等； < 0.40 為差 (Salter et al., 2005)。

收斂效度

收斂效度皮爾森相關係數 (r) 分別分析 SIS-ML 各向度之總分與效標測驗分數之相關程度。收斂效度之效標測驗包含 NIHSS 及 BI。然而，SIS 包含生理、心理、社會及認知等向度，因此，SIS 之肌力、活動力、日常生活自理能力及手部功能 4 個生理相關之向度與 NIHSS 及 BI 所評估之向度相關較高。之前研究驗證 SIS 各向度與 NIHSS 及 BI 之收斂效度彙整於表三。因 SIS 之收斂效度標準依各向度有所差別，NIHSS 及 BI 於肌力、活動力、日常生活自理能力及手部功能向度之 $r \geq 0.50$ 為良好； $0.20-0.49$ 為中等； < 0.20 為差；社會參與能力之 $r \geq 0.40$ 為良好； $0.30-0.39$ 為中等； < 0.30 為差；情緒、溝通能力、記憶力之 $r \geq 0.20$ 為良好； $0.10-0.19$ 為中等； < 0.10 為差 (Carod-Artal et al., 2008; Chou et al., 2015)。



第六章 研究結果



第一節 第一階段：SIS-ML 之發展

步驟 1：挑選短版題目之題組

研究者使用第一層隱藏層包含套索迴歸之隱藏層 4 層，人工神經元個數 32 個之 ANN 模型分析各題對於 SIS 題庫之預測性，並以第一層隱藏層之模型參數絕對值加總值作為各題預測性指標，以作為選題之標準。選題之 ANN 模型圖如圖七所示，各題之模型參數絕對值加總值列於圖八。最終研究者篩選出之 17 組短版題組。各向度之後選題目列於附錄一，各題組之題目項目列於附錄二。

步驟 2：使用不同模型架構訓練以短版題目模擬原版分數之 ANN 模型

研究者將訓練資料匯入 ANN 模型中，17 組題目分別有不同之 ANN 模型以獲得其模擬分數，最終獲得 136 個 ANN 模型及模擬分數。結果顯示於 6X196、10X196 及 6X512 模型中，所有題組於 8 個向度都有良好之模型分數 ($R^2 = 0.98-0.99$)。於 10X512 模型中，僅 16~20、23~26、30、32 題題組之 8 個向度模型分數皆達良好 ($R^2 = 0.80-0.99$)。於 10X32 模型中，僅 18、20、30 題題組之 8 個向度模型分數皆達良好 ($R^2 = 0.88-0.99$)。於 6X8、10X8、6X32 模型中，所有題組僅部分向度達到可接受至良好之模型分數 ($R^2 = 0.60-0.99$)。各模型之訓練資料之模型分數詳見表四、表五、表六、表七。

步驟 3：使用訓練資料挑選預測力較佳之 ANN 模型架構

研究者依據步驟二呈現之模型分數，挑選適合之 ANN 模型架構。結果顯示 17 組題組之模型分數皆達良好之 ANN 模型架構為 6X196、10X196 及 6X512。各 ANN 模型於各題組之預測力列於圖九。

步驟 4：使用驗證資料挑選預測力較佳之題數組合

研究者保留步驟三中 8 個向度之模型分數都達良好之題組，並匯入驗證資料以進行交叉驗證。各題組之驗證資料模型分數如表八。結果顯示 27 至 32 題於上述 6X196、10X196、6X512 三個 ANN 模型架構中，8 個向度之驗證資料之模型

分數仍達良好($R^2 = 0.80-0.99$)。16 至 26 題題組中，僅達可接受至良好之模型分數 ($R^2 = 0.60-0.99$)。



步驟 5：使用測試資料選擇最適之機器學習模型

研究者於步驟四驗證模型後，匯入測試資料，並確認模型是否於測試資料仍有良好之模型分數。候選之 6 組題組之訓練、驗證及測試資料模型分數如表九。結果顯示 27 至 32 題題組於訓練、驗證及測試資料於 3 個 ANN 模型架構中之 8 個向度之測試資料之模型分數仍達良好 ($R^2 = 0.80-0.99$)。研究者使用 6 組候選題組之訓練、驗證及測試資料之模型分數平均作為選擇之依據，最終選擇 6X196 之 27 題題組之版本為 SIS-ML 之題組及模擬分數以進行研究二之心理計量特性驗證。6 組候選題之模型分數平均列於表十。SIS-ML 於測試資料之記憶力、情緒、溝通能力、日常生活自理能力及活動力向度分數有可接受之差異百分比 (4% - 10%)，於肌力、手部功能及社會參與能力有較大之差異百分比 (11-16%)。測試資料之最終版 SIS-ML、原版 SIS 及兩者差異之平均分數、標準差、差異百分比列於表十一。27 題題組於 6X196 模型之模擬分數與原始分數之散佈圖如圖十，題組及 ANN 模型圖如圖十一。

第二節 第二階段：驗證 SIS-ML 之心理計量特性

研究者將測試資料之 SIS-ML 分數與原版 SIS 分數、NIHSS 及 BI 之皮爾森相關係數作為 SIS-ML 之同時效度與收斂效度。結果如表十二所示。

同時效度部分，結果顯示 SIS-ML 分數與 SIS 分數達良好之同時效度 ($r = 0.92-0.99$)。收斂效度部分，結果顯示 SIS-ML 分數與 NIHSS 於肌力、日常生活自理能力、活動力及手部功能達良好之收斂效度 ($r = -0.54- -0.59$)，於社會參與能力達良好之收斂效度 ($r = -0.58$)，於情緒、溝通能力、記憶力達良好之收斂效度 ($r = -0.34- -0.42$)。SIS-ML 分數與 BI 於日常生活自理及活動力達良好之收斂效度 ($r = 0.56-0.74$)，於肌力及手部功能僅達中度之收斂效度 ($r = 0.41-0.43$)，於社會參與能力達良好之收斂效度 ($r = 0.44$)，於情緒、記憶力及溝通能力皆達良好之收斂效度 ($r = 0.22-0.25$)。

第七章 討論



本研究目的為運用 ML 模型之預測力，以發展 SIS-ML。此外，研究者亦初步驗證 SIS-ML 之同時效度及收斂效度。

短版測驗之題組

本研究測試 16 題至 32 題共 17 組題組，找出最適合之短版題目為 27 題之題組。結果顯示所有題組在選定適合之 ANN 模型架構時，皆可於訓練資料達良好之模型分數。然而，當進行驗證資料之交叉驗證時，僅 27 至 32 題之題組於訓練資料及驗證資料達良好之模型分數，代表 27 題至 32 題題目所訓練出來之模型可類化到其他資料，且不會產生 ANN 模型過度擬合當前資料之問題。6 組題組之中，27 題為題目最少且模型之模擬分數仍可達到良好之短版題組。16 題至 26 題之題組於交叉驗證中，情緒向度之題目預測力僅能達到差至可接受之模型分數，代表自 26 題以下之題組題目訓練出來之 ANN 模型可能會出現過度擬合當前資料之問題，所訓練之預測模型無法類化至其它資料。

效度驗證

同時效度檢驗結果顯示測試資料之 SIS-ML 與原版 SIS 各向度總分之相關係數皆達 0.9，初步支持 SIS-ML 及原版 SIS 達良好之同時效度。SIS-ML 可縮短原版 SIS 之評估題數，並可藉由訓練好之 ANN 模型模擬 SIS 分數，因此 SIS-ML 可保有原版 SIS 所評估之 8 個向度。本研究結果亦支持使用 ML 模型發展之 SIS-ML 可評估到與原版 SIS 相同之向度。然而，由 SIS-ML 分數及原版 SIS 分數散佈圖可看出部分個案之 SIS-ML 之分數仍有較大之誤差，極端值之預測力亦有較大之誤差。此外，SIS-ML 於肌力、手部功能及社會參與能力向度與原版 SIS 之差異百分比比較其他向度大，代表 SIS-ML 於肌力、手部功能及社會參與能力之預測精準度與其他向度相比，仍顯不足。因此，後續研究者使用 SIS-ML 時，若個案之評估結果較差時，可考慮施測所有項目以獲得較精準之數據，以確認個案之 HRQOL。

收斂效度的部分，SIS-ML 與 NIHSS 及 BI 之相關程度與過去研究之原版 SIS



於 BI 之相關程度一致 (Carod-Artal et al., 2008; Choi et al., 2017; Chou et al., 2015; Vellone et al., 2015)。SIS-ML 與 NIHSS 於肌力、日常生活自理能力、活動力及手部功能及社會參與能力之相關係數皆達 0.5，於情緒、溝通能力、記憶力皆達 0.3，此數值符合過去文獻所示之健康生活品質與整體症狀之關聯程度 (Carod-Artal et al., 2008; Choi et al., 2017; Chou et al., 2015)。本研究所發展之 SIS-ML 於情緒、溝通能力及記憶力向度之相關程度 ($r = -0.34 - 0.42$) 與過去研究之原版 SIS 於此 3 個向度之相關程度 ($r = -0.08 - 0.30$) 一致，皆呈現低度相關 (Carod-Artal et al., 2008; Choi et al., 2017; Chou et al., 2015; Vellone et al., 2015)。SIS-ML 與 BI 於日常生活自理及活動力之相關係數皆達 0.60，於肌力、手部功能及社會參與能力之相關係數達 0.40，於情緒、記憶力及溝通能力皆之相關係數達 0.22。過去研究中，原版 SIS 之日常生活自理能力及活動力與 BI 呈高度相關，於肌力、手部功能及社會參與能力呈中度相關，於情緒、記憶力及溝通能力呈低度相關 (Carod-Artal et al., 2008; Choi et al., 2017)。總體而言，SIS-ML 各向度評估之能力符合研究者預期，即 SIS-ML 可評估原版 SIS 所評估之 8 個向度。

SIS-ML 之實務應用價值

本研究初步支持 SIS-ML 可縮短原版 SIS 之題數至少一半之題目，SIS-ML 可有效縮短評估個案 HRQOL 所耗費之時間，有助於縮短個案填寫之時間。且 SIS-ML 可保有原版 SIS 之評估向度，可完整評估個體各向度之 HRQOL，以利臨床人員掌握個案於復健進程中之 HRQOL。然而，SIS-ML 仍為經過機器學習模型模擬之分數，無法達到 100% 預測原版 SIS 之能力，且 SIS-ML 於肌力、手部功能及社會參與能力與原版 SIS 有較大之誤差。後續臨床人員於使用該工具時，宜參考 SIS-ML 與原版 SIS 之標準差及差異百分比，尤其個案呈現較差分數時，可考慮填寫所有項目，以獲得較精準之數據。

本研究所發展之 SIS-ML 需要轉換器或轉換平台以模擬原版 SIS 之分數。目前研究者預計建構電腦化之 SIS-ML，可供個案、研究人員或臨床人員快速評估 SIS-ML，並可快速轉換 SIS-ML 之分數。平台預計以網頁呈現，以利個案、研究人員或臨床人員使用，提升評估效能。



ML 用於發展短版測驗之預測效能

本研究結果初步支持 ML 技術可讓 SIS 縮短至 27 題，且所獲得之模擬分數可逼近與原版 SIS 所得之結果。研究者使用 6 層及 10 層隱藏層及人工神經元數為 8、32、196、512 之機器學習模型以訓練短版題目模擬原版分數之 ANN 模型。結果顯示 ML 技術可保有原版題目包含 8 個向度之架構，且此預測模型亦包含各題跨向度之預測能力。然而，結果顯示於本研究中，6X196、10X196 及 6X512 之 ML 模型為候選模型中相對良好之模型架構。此結果也顯示當模型架構過於簡單（6X8、10X8、6X32、10X32）或過於複雜時（10X512），模型之預測能力不佳。由此可知，機器學習模型之複雜程度與模型之預測能力並非呈現正相關。

本研究所使用之研究方法為嘗試多種版本之 SIS-ML 題組，並選擇整體機器學習模型最佳之題組作為 SIS-ML 之題目。結果顯示 SIS-ML 之分數與原版 SIS 分數之誤差達可接受之範圍。然而，仍有許多題組及 ANN 模型配對之可能性尚未驗證，且 SIS-ML 於肌力、手部功能及社會參與能力之分數與原版 SIS 分數仍有較大之誤差。後續研究可使用不同選題方式選出預測效能更佳之題組，以降低 SIS-ML 預測分數之誤差。

研究限制

本研究所發展之 SIS-ML 具有下列四項弱點，可能降低其應用價值：(1) 樣本代表性不足：本研究所使用之資料為方便樣本，且僅招募慢性之中風患者。故本研究結果恐難以類推至急性或亞急性中風患者。再者，本研究所使用之資料為個案填寫之版本，亦難類推至臨床人員或家屬代填之版本。(2) 樣本數較少：本研究所使用之資料僅 256 位患者。目前尚未有文獻指出適用於獲得精準 ANN 模型之參數估計值所需之樣本數，然而，過去機器學習模型之文獻多使用樣本數為 1000 以上之樣本數 (Heo et al., 2019; Senders et al., 2018; Tozlu et al., 2020)。故本研究所訓練之機器學習模型可能受限於樣本數之不足，而使模型之預測能力較不準確。(3) 尚未進行臨床試驗：SIS-ML 目前僅使用過去收案之資料進行模擬分析，尚未使用 SIS-ML 進行臨床試驗。未來研究需再使用獨立之樣本驗證 SIS-ML 之信度與效度。



未來研究建議

未來研究包含二大方向：(1) 使用更大之樣本數確認目前版本及目前模型之預測力；(2) SIS-ML 之後續驗證。本研究初步支持 ML 模型可發展短版之測驗，然而，未來應使用更大之樣本數以確認目前版本為最適之 SIS-ML 及最適之 ML 模型。

目前之 SIS-ML 後續待驗證之項目至少有二：(1) 增加研究樣本數：目前版本之 SIS-ML 尚未進行臨床試驗，因此未來須持續收集 SIS 之資料，並驗證目前版本之 SIS-ML 之預測力及信、效度，以確認其臨床適用性。(2) 擴展 SIS-ML 之適用對象：由於急性或亞急性個案亦存在 HRQOL 之障礙，故未來應針對不同病程之個案驗證 SIS-ML 之心理計量特性，以確認目前版本之 SIS 亦可適用於不同病程之族群。

第八章 總結



本研究目的為以發展 SIS-ML 及初步驗證其同時效度及收斂效度。研究者依據機器學習模型之預測能力，發展 SIS-ML 之題組，並驗證 SIS-ML 之模擬分數與原始分數之一致性，以確認該模型可類化至其他資料。確認題組及模型後，研究者再驗證 SIS-ML 之同時效度及收斂效度。

第一階段之研究結果顯示 27 題之題組於 6 層隱藏層及 196 個人工神經元之模型有最好之模型分數，代表該模型之 SIS-ML 模擬分數與原版 SIS 分數有最好之一致性。研究結果亦初步支持機器學習模型之預測力可用於發展短版測驗，以提升評估工具之評估效能。第二階段之研究結果支持 27 題題組之 SIS-ML 與原版 SIS 有良好之同時效度，與 NIHSS 及 BI 亦有良好之收斂效度，代表 SIS-ML 可保有原版 SIS 之 8 個向度。

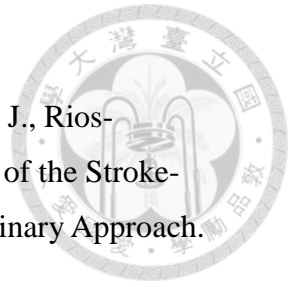
然而，本研究所使用之樣本數較少，且本研究發展之 SIS-ML 尚未進行臨床驗證，因此未來研究需使用更大之樣本數確認目前版本及目前模型之預測力及再使用獨立之樣本驗證 SIS-ML 之信效度。

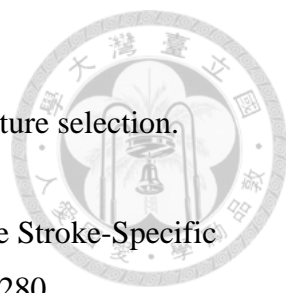
參考資料

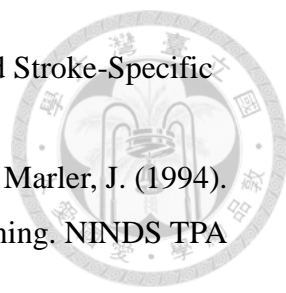


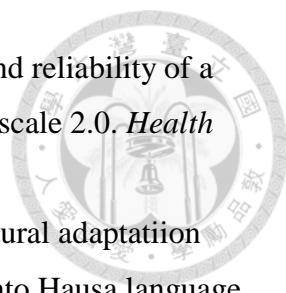
- 台灣版世界衛生組織生活品質問卷發展小組. (2000). 台灣版世界衛生組織生活品質問卷之發展簡介. *中華公共衛生雜誌*, *19*, 315-324.
- 姚開屏. (2002). 健康相關生活品質概念與測量原理之簡介. *台灣醫學*, *6*, 183-192.
- 梁佩蓉, & 林佩欣. (2016). 中風衝擊量表用於台灣中風患者之信效度測量. *物理治療*, *41*, 28-36.
- Alhamzawi, R., & Ali, H. T. M. (2018). The Bayesian adaptive lasso regression. *Mathematical Biosciences*, *303*, 75-82.
- Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports medicine*, *26*, 217-238.
- Bleidorn, W., & Hopwood, C. J. (2019). Using Machine Learning to Advance Personality Assessment and Theory. *Personality and Social Psychology Review*, *23*, 190-203.
- Carod-Artal, F. J., Coral, L. F., Trizotto, D. S., & Moreira, C. M. (2008). The stroke impact scale 3.0: evaluation of acceptability, reliability, and validity of the Brazilian version. *Stroke*, *39*, 2477-2484.
- Chang, K. C., & Tseng, M. C. (2003). Costs of acute care of first-ever ischemic stroke in Taiwan. *Stroke*, *34*, e219-221.
- Chen, H. F., Wu, C. Y., Lin, K. C., Li, M. W., & Yu, H. W. (2012). Validity, reliability and responsiveness of a short version of the Stroke-Specific Quality of Life Scale in patients receiving rehabilitation. *Journal of rehabilitation medicine*, *44*, 629-636.
- Choi, S. U., Lee, H. S., Shin, J. H., Ho, S. H., Koo, M. J., Park, K. H., . . . Kim, D. A. (2017). Stroke Impact Scale 3.0: Reliability and Validity Evaluation of the Korean Version. *Annals of Translational Medicine*, *41*, 387-393.
- Chou, C. Y., Ou, Y. C., & Chiang, T. R. (2015). Psychometric comparisons of four disease-specific health-related quality of life measures for stroke survivors. *Clinical Rehabilitation*, *29*, 816-829.
- Collin, C., Wade, D. T., Davies, S., & Horne, V. (1988). The Barthel ADL Index: a

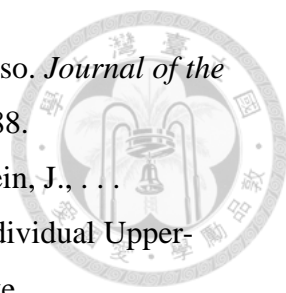
- reliability study. *International disability studies*, 10, 61-63.
- Cruz-Cruz, C., Martinez-Nunez, J. M., Perez, M. E., Kravzov-Jinich, J., Rios-Castaneda, C., & Altagracia-Martinez, M. (2013). Evaluation of the Stroke-Specific Quality-of-Life (SSQOL) Scale in Mexico: A Preliminary Approach. *Value in Health Regional Issues*, 2, 392-397.
- Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, 132, 1920-1930.
- Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *The Journal of Nervous and Mental Disease*, 161, 185-190.
- Duncan, P. W., Lai, S. M., Bode, R. K., Perera, S., & DeRosa, J. (2003). Stroke Impact Scale-16: A brief assessment of physical function. *Neurology*, 60, 291-296.
- Duncan, P. W., Wallace, D., Lai, S. M., Johnson, D., Embretson, S., & Laster, L. J. (1999). The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke*, 30, 2131-2140.
- Emsley, H. C., & Hopkins, S. J. (2008). Acute ischaemic stroke and infection: recent and emerging concepts. *The Lancet Neurology*, 7, 341-353.
- Fonti, V., & Belitser, E. (2017). Feature selection using lasso. *VU Amsterdam Research Paper in Business Analytics*, 1-25.
- Freedman, D., Pisani, R., & Purves, R. (2007). *Statistics: Fourth International Student Edition*: W.W. Norton & Company.
- Ganvir, S., Harishch, M., & Kunde, C. (2018). Validation of Marathi Version of Stroke-specific Quality of Life. *Annals of Indian Academy of Neurology*, 21, 290-293.
- Gazzaniga, M. S. (2014). *Cognitive neuroscience : the biology of the mind / Michael S. Gazzaniga, University of California, Santa Barbara; Richard B. Ivry, University of California, Berkeley; George R. Mangun, University of California, Davis* (Fourth edition. ed.). New York: W. W. Norton & Company, Inc.
- Geyh, S., Cieza, A., & Stucki, G. (2009). Evaluation of the German translation of the Stroke Impact Scale using Rasch analysis. *Clinical Neuropsychologist*, 23, 978-995.
- Giger, M. L. (2018). Machine Learning in Medical Imaging. *Journal of the American College of Radiology*, 15, 512-520.
- Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality

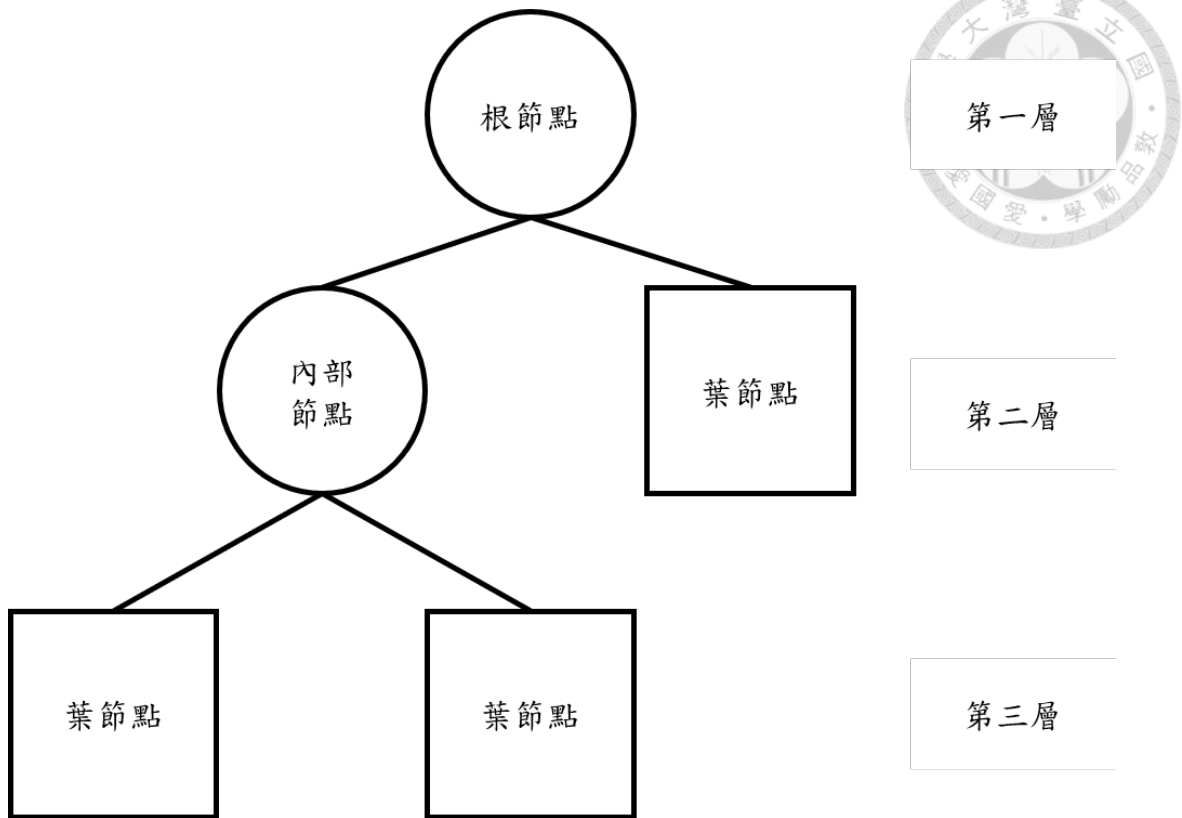


- 
- of life. *Annals of Internal Medicine*, 118, 622-629.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3, 1157-1182.
- Hakverdioglu Yont, G., & Khorshid, L. (2012). Turkish version of the Stroke-Specific Quality of Life Scale. *International Nursing Review*, 59, 274-280.
- Heo, J., Yoon, J. G., Park, H., Kim, Y. D., Nam, H. S., & Heo, J. H. (2019). Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. *Stroke*, 50, 1263-1265.
- Hobart, J. C., Lamping, D. L., & Thompson, A. J. (1996). Evaluating neurological outcome measures: the bare essentials. *Journal of neurology, neurosurgery, and psychiatry*, 60, 127-130.
- Hsueh, I. P., Jeng, J. S., Lee, Y., Sheu, C. F., & Hsieh, C. L. (2011). Construct validity of the stroke-specific quality of life questionnaire in ischemic stroke patients. *Archives of Physical Medicine and Rehabilitation*, 92, 1113-1118.
- Hu, H. H., Chu, F. L., Chiang, B. N., Lan, C. F., Sheng, W. Y., Lo, Y. K., . . . Luk, Y. O. (1989). Prevalence of stroke in Taiwan. *Stroke*, 20, 858-863.
- Hu, H. H., Sheng, W. Y., Chu, F. L., Lan, C. F., & Chiang, B. N. (1992). Incidence of stroke in Taiwan. *Stroke*, 23, 1237-1241.
- Jain, A. K., Mao, J. C., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29, 31-+.
- Kamwesiga, J. T., von Koch, L., Kottorp, A., & Guidetti, S. (2016). Cultural adaptation and validation of Stroke Impact Scale 3.0 version in Uganda: A small-scale study. *SAGE Open Medicine*, 4, 2050312116671859.
- Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of chronic diseases*, 38, 27-36.
- Kline, P. (1998). *The new psychometrics: Science, psychology, and measurement*: Psychology Press.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436-444.
- Legris, N., Devilliers, H., Dumas, A., Carnet, D., Charpy, J. P., Bastable, P., . . . Bejot, Y. (2018). French validation of the Stroke Specific Quality of Life Scale (SS-QoL). *NeuroRehabilitation*, 42, 17-27.
- Lin, K. C., Fu, T., Wu, C. Y., Hsieh, Y. W., Chen, C. L., & Lee, P. C. (2010).

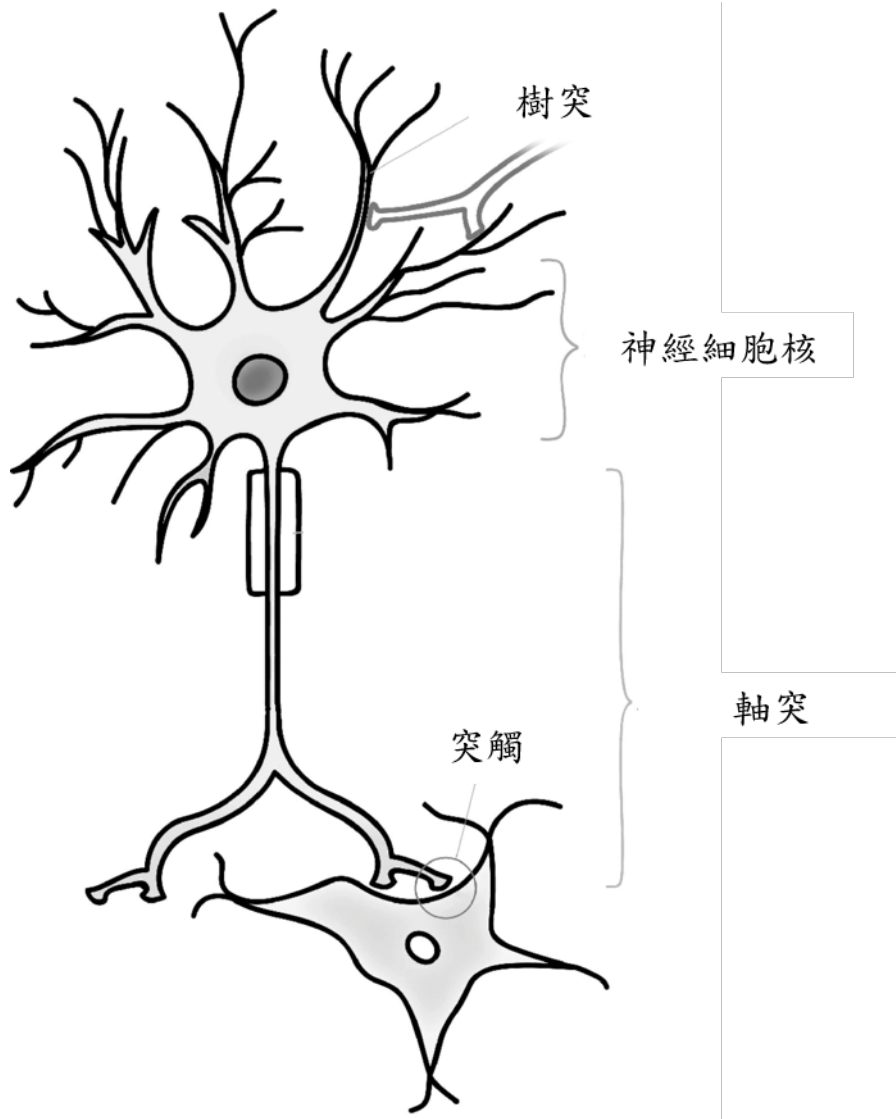
- 
- Psychometric comparisons of the Stroke Impact Scale 3.0 and Stroke-Specific Quality of Life Scale. *Quality of Life Research*, 19, 435-443.
- Lyden, P., Brott, T., Tilley, B., Welch, K., Mascha, E., Levine, S., . . . Marler, J. (1994). Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke*, 25, 2220-2226.
- MacIsaac, R., Ali, M., Peters, M., English, C., Rodgers, H., Jenkinson, C., . . . Collaboration, V. (2016). Derivation and Validation of a Modified Short Form of the Stroke Impact Scale. *Journal of the American Heart Association*, 5.
- Mackay, J., & Mensah, G. A. (2004). *The atlas of heart disease and stroke*: World Health Organization.
- Mahmoodi, M., Safari, A., Vossoughi, M., Golbon-Haghighi, F., Kamali-Sarvestani, M., Ghaem, H., & Borhani-Haghighi, A. (2015). Stroke specific quality of life questionnaire: Test of reliability and validity of the Persian version. *Iranian Journal of Neurology*, 14, 94-100.
- Mohammad, A. H., Al-Sadat, N., Siew Yim, L., & Chinna, K. (2014). Reliability and validity of the Nigerian (Hausa) version of the Stroke Impact Scale (SIS) 3.0 index. *BioMed Research International*, 2014, 302097.
- Mokkink, L. B., Prinsen, C. A., Bouter, L. M., Vet, H. C., & Terwee, C. B. (2016). The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian journal of physical therapy*, 20, 105-113.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, 63, 737-745.
- Muus, I., Williams, L. S., & Ringsberg, K. C. (2007). Validation of the Stroke Specific Quality of Life Scale (SS-QOL): test of reliability and validity of the Danish version (SS-QOL-DK). *Clinical Rehabilitation*, 21, 620-627.
- Ochi, M., Ohashi, H., Hachisuka, K., & Saeki, S. (2017). The Reliability and Validity of the Japanese Version of the Stroke Impact Scale Version 3.0. *Journal of UOEH*, 39, 215-221.

- 
- Odetunde, M. O., Akinpelu, A. O., & Odole, A. C. (2017). Validity and reliability of a Nigerian-Yoruba version of the stroke-specific quality of life scale 2.0. *Health and Quality of Life Outcomes, 15*, 205.
- Odetunde, M. O., Akinpelu, A. O., & Odole, A. C. (2018). Cross-cultural adaptation and validation of the stroke specific quality of life 2.0 scale into Hausa language. *Journal of Patient-Reported Outcomes, 2*, 63.
- Pedersen, S. G., Heiberg, G. A., Nielsen, J. F., Friberg, O., Stabel, H. H., Anke, A., & Arntzen, C. (2018). Validity, reliability and Norwegian adaptation of the Stroke-Specific Quality of Life (SS-QOL) scale. *SAGE Open Medicine, 6*, 2050312117752031.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of machine learning research, 12*, 2825-2830.
- Pinter, M. M., & Brainin, M. (2012). Rehabilitation after stroke in older people. *Maturitas, 71*, 104-108.
- Portney, L. G. (2020). *Foundations of Clinical Research: Applications to Evidence-Based Practice*: FA Davis.
- Sallam, S. A., Al-Khamis, F. A., Muaidi, Q. I., & Abdulla, F. A. (2019). Translation and validation of the stroke specific quality of life scale into Arabic. *NeuroRehabilitation, 44*, 283-293.
- Salter, K., Jutai, J. W., Teasell, R., Foley, N. C., Bitensky, J., & Bayley, M. (2005). Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disability and Rehabilitation, 27*, 315-340.
- Sbordone, R. J., & Long, C. (1996). *Ecological validity of neuropsychological testing*: CRC Press.
- Senders, J. T., Staples, P. C., Karhade, A. V., Zaki, M. M., Gormley, W. B., Broekman, M. L. D., . . . Arnaout, O. (2018). Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurgery, 109*, 476-486 e471.
- Talabis, M. R. M., McPherson, R., Miyamoto, I., & Martin, J. L. (2014). *Information Security Analytics: Finding Security Insights, Patterns, and Anomalies in Big Data*: Syngress Publishing.

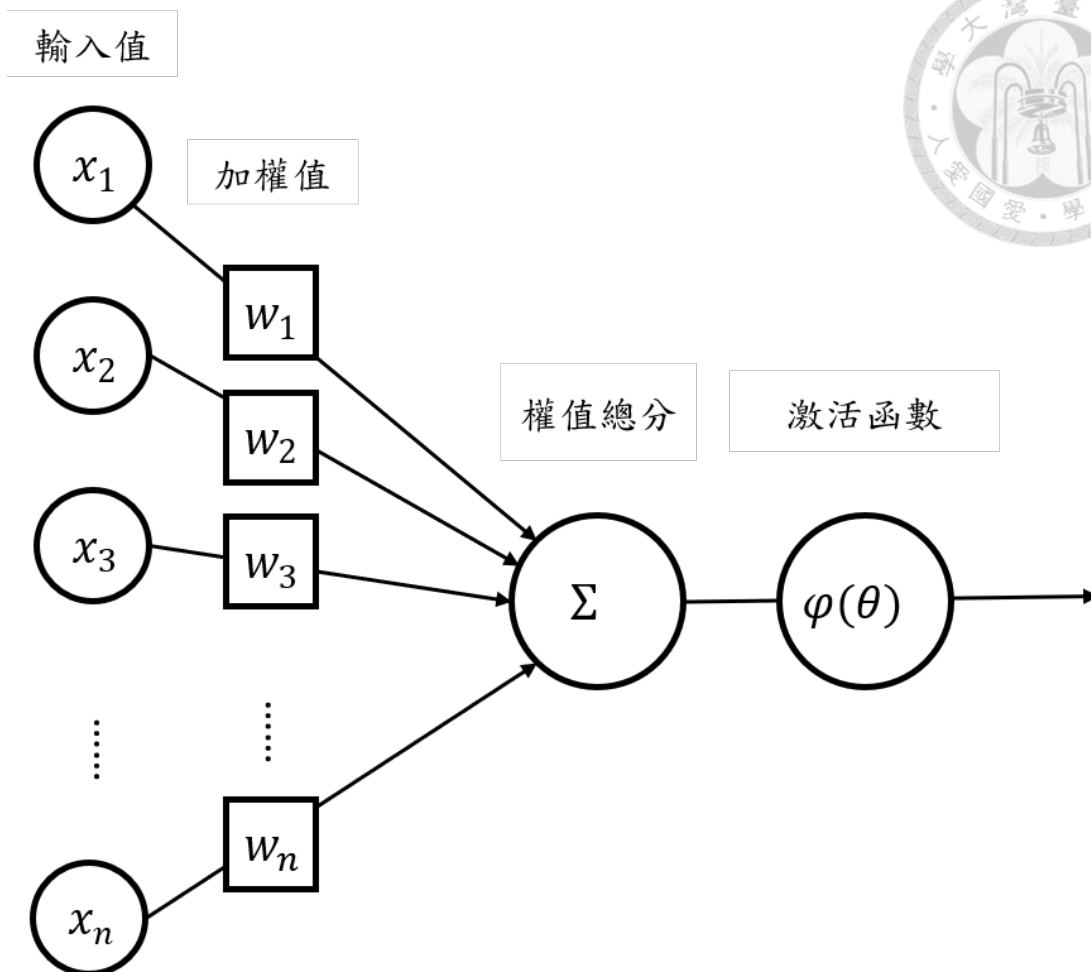
- 
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society Series B-Methodological*, 58, 267-288.
- Tozlu, C., Edwards, D., Boes, A., Labar, D., Tsagaris, K. Z., Silverstein, J., . . . Kuceyeski, A. (2020). Machine Learning Methods Predict Individual Upper-Limb Motor Impairment Following Therapy in Chronic Stroke. *Neurorehabilitation and Neural Repair*, 34, 428-439.
- Truelsen, T., Begg, S., & Mathers, C. (2006). *The global burden of cerebrovascular disease*. Paper presented at the Who Int.
- Van Merriënboer, B., Bahdanau, D., Dumoulin, V., Serdyuk, D., Warde-Farley, D., Chorowski, J., & Bengio, Y. (2015). Blocks and fuel: Frameworks for deep learning. *arXiv preprint arXiv:1506.00619*.
- Vellone, E., Savini, S., Fida, R., Dickson, V. V., Melkus, G. D., Carod-Artal, F. J., . . . Alvaro, R. (2015). Psychometric evaluation of the Stroke Impact Scale 3.0. *Journal of Cardiovascular Nursing*, 30, 229-241.
- Williams, L. S., Weinberger, M., Harris, L. E., Clark, D. O., & Biller, J. (1999). Development of a stroke-specific quality of life scale. *Stroke*, 30, 1362-1369.
- World-Health-Organisation. (1978). *Cerebrovascular disorders*. Geneva: WHO.
- Yin, S., Njai, R., Barker, L., Siegel, P. Z., & Liao, Y. (2016). Summarizing health-related quality of life (HRQOL): development and testing of a one-factor model. *Population health metrics*, 14, 22.
- Zhang, Z. (2016). A gentle introduction to artificial neural networks. *Annals of Translational Medicine*, 4, 370.



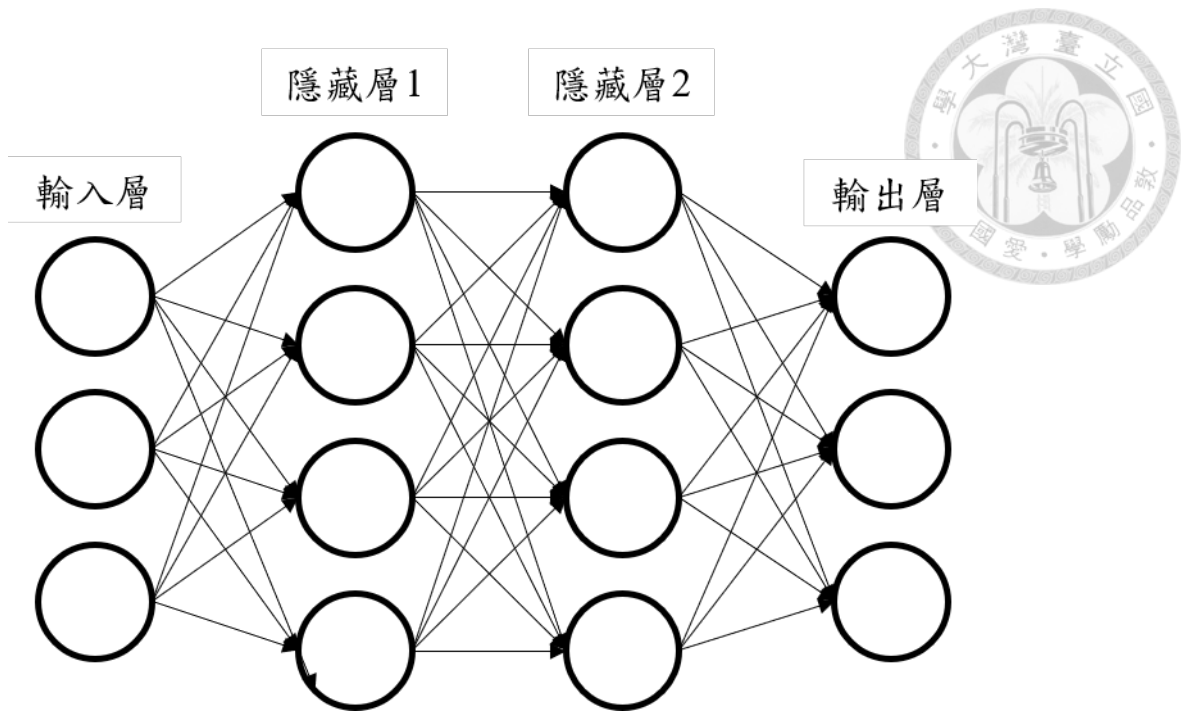
圖一：決策樹之基本結構



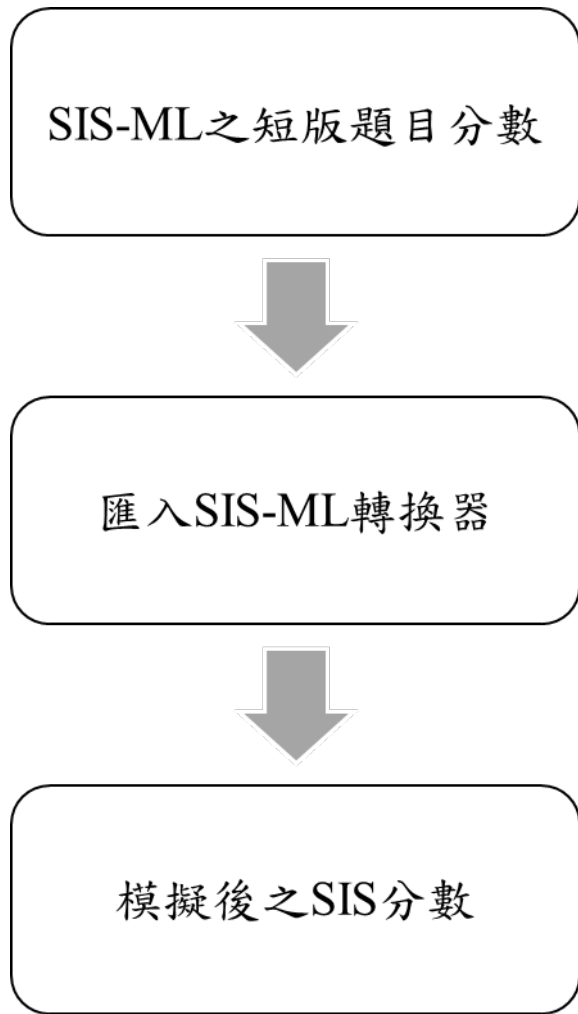
圖二：腦神經細胞



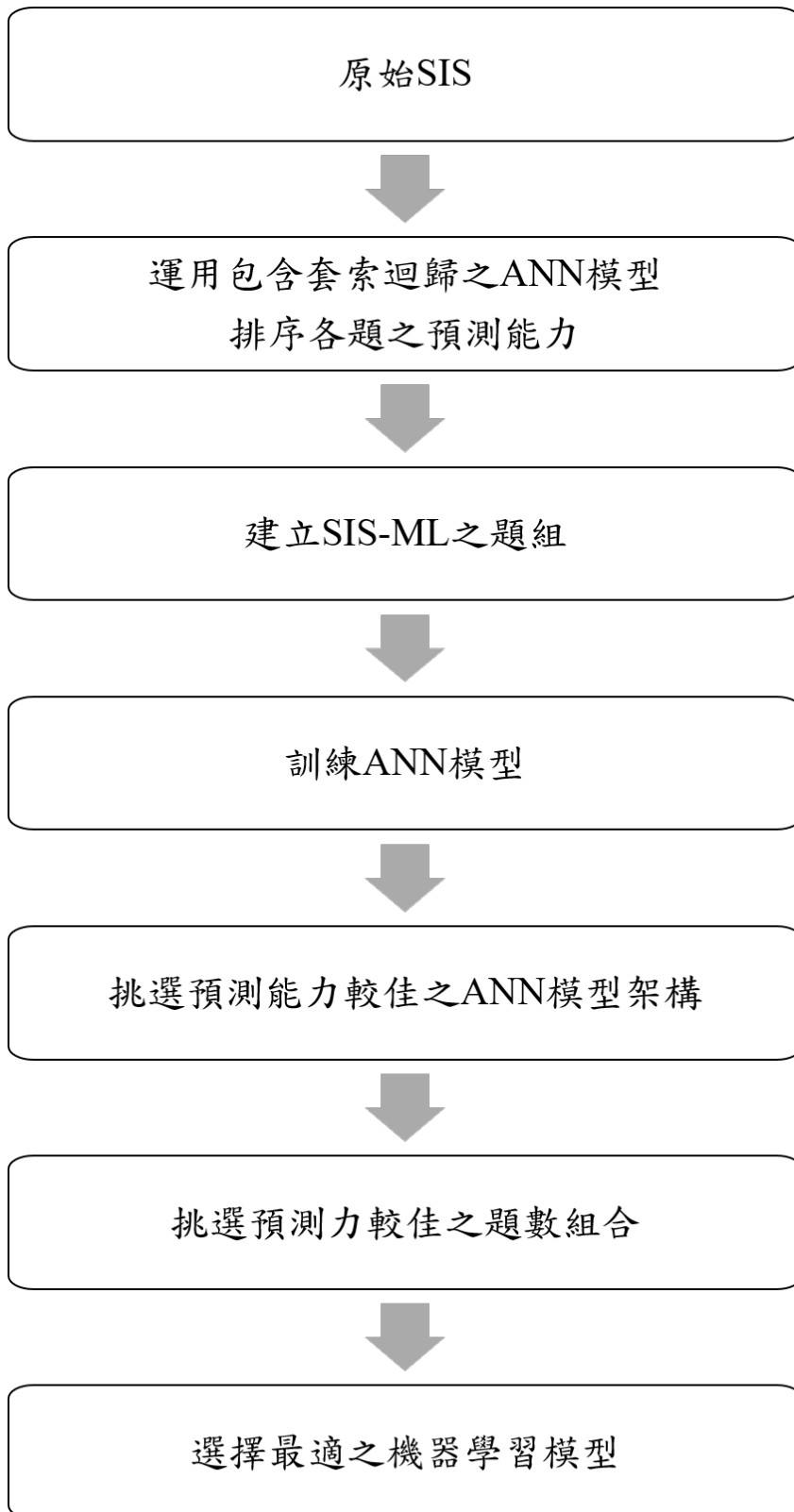
圖三：人工神經元



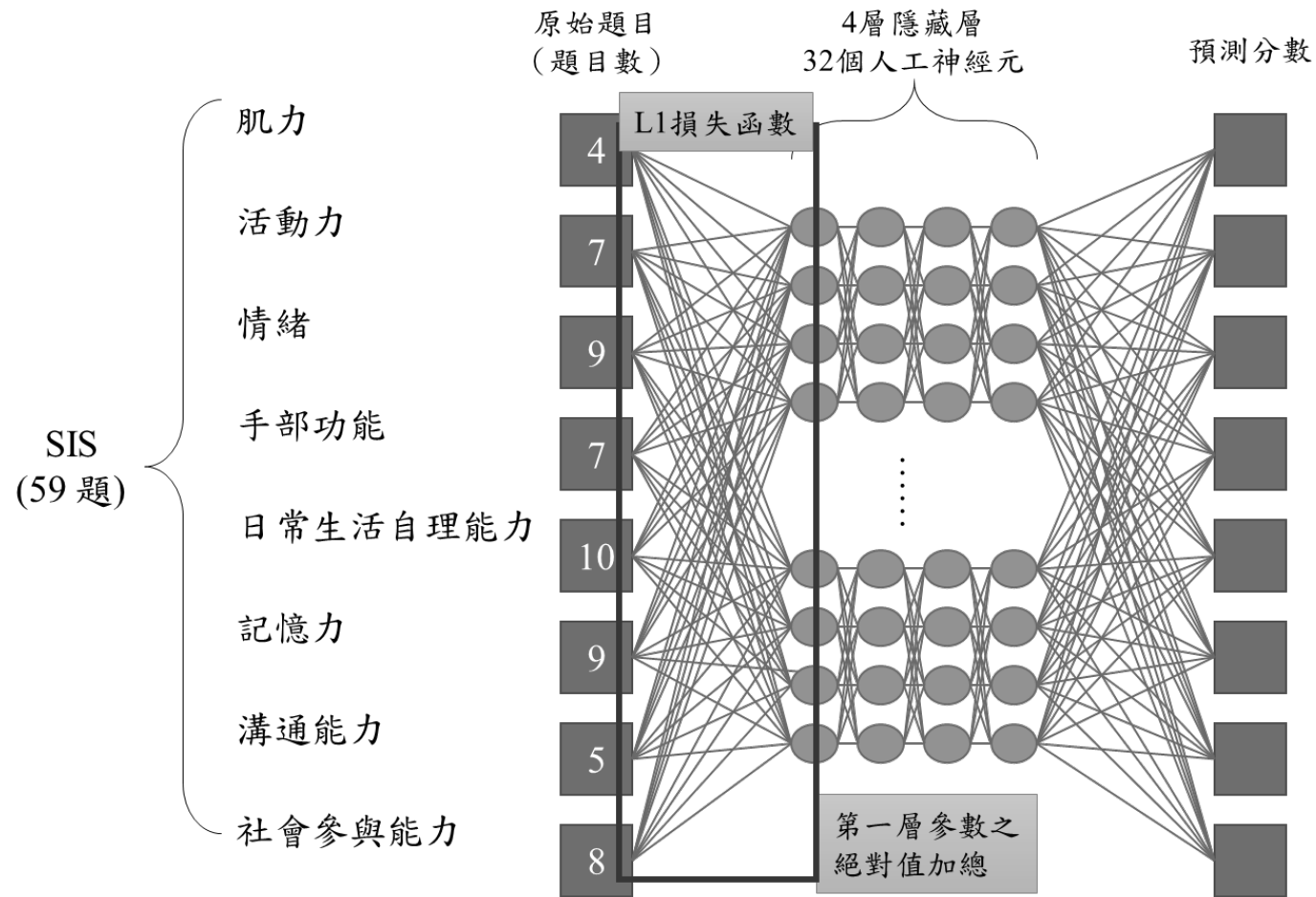
圖四：人工神經網路（以 2 層隱藏層為例）



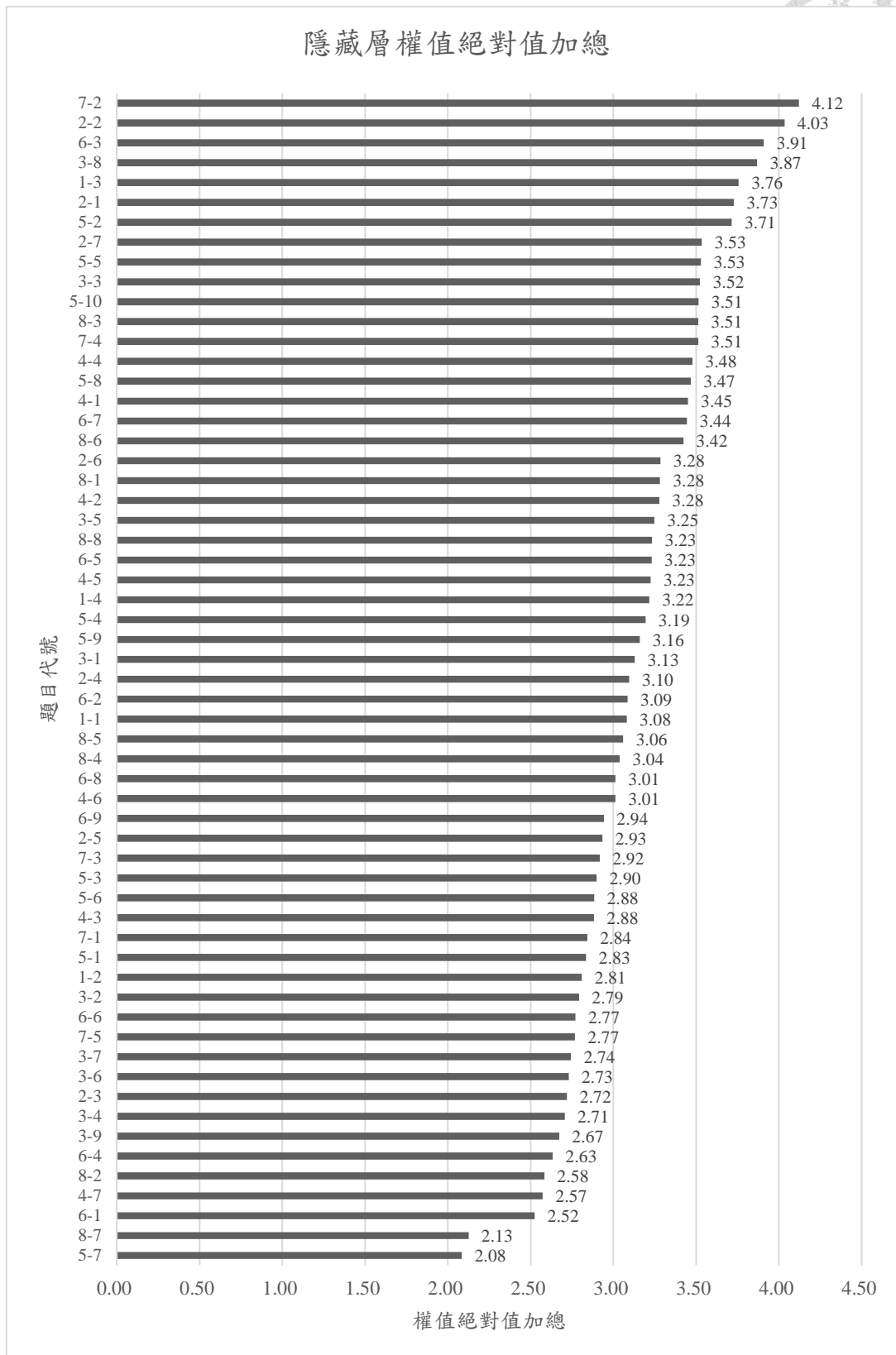
圖五：SIS-ML 之分數轉換流程



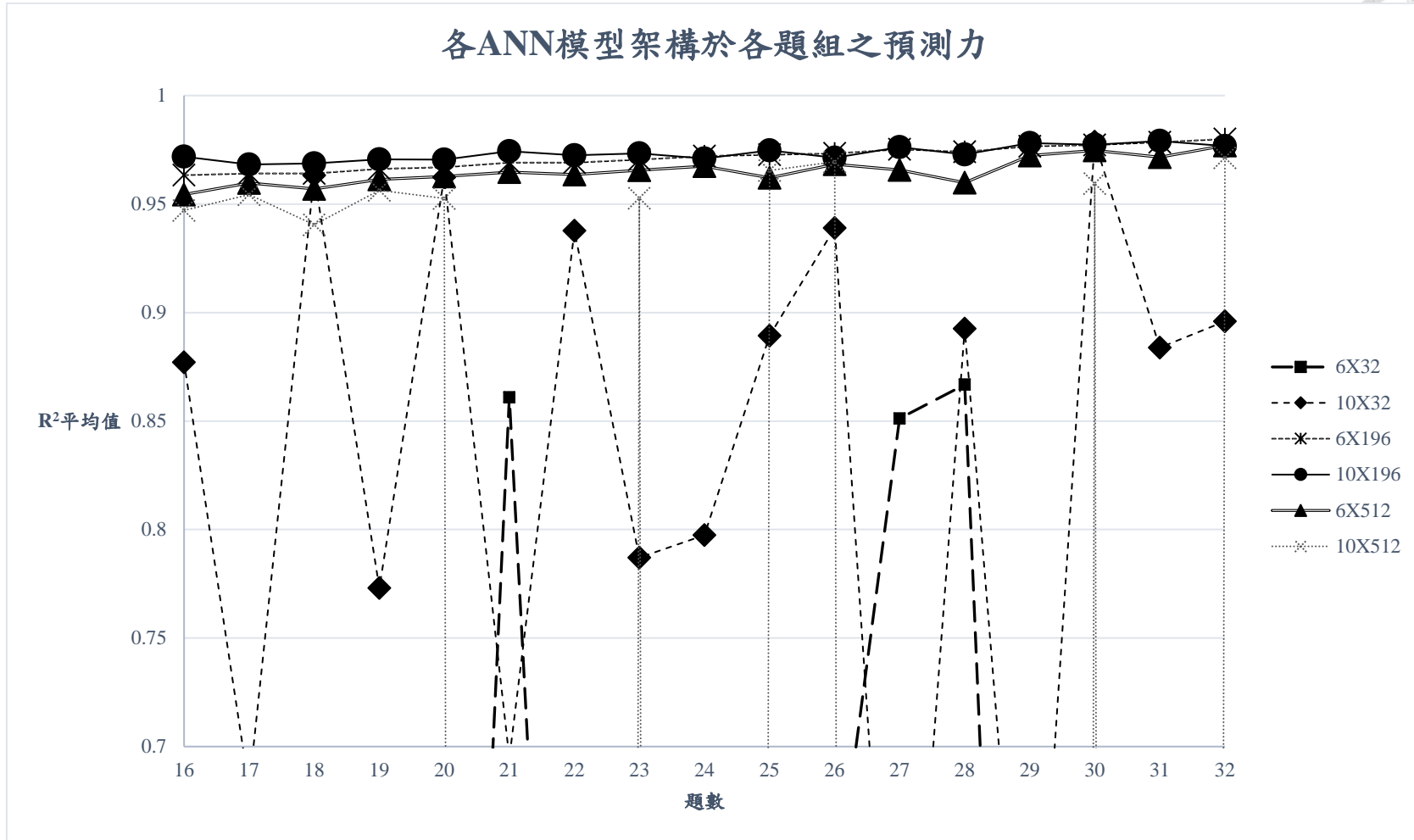
圖六：SIS-ML 發展流程



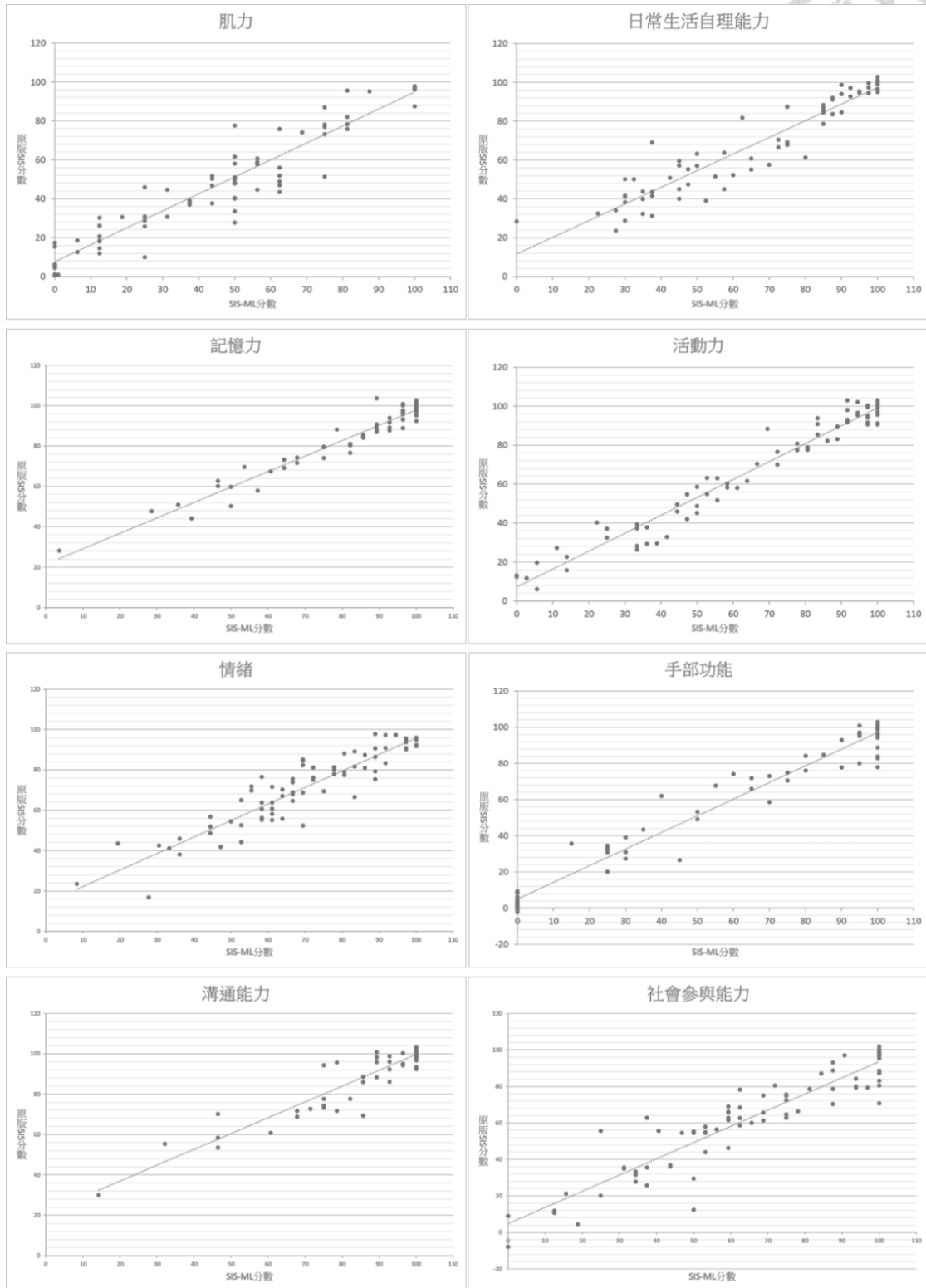
圖七：步驟 1 所使用之包含套索迴歸之 ANN 模型圖



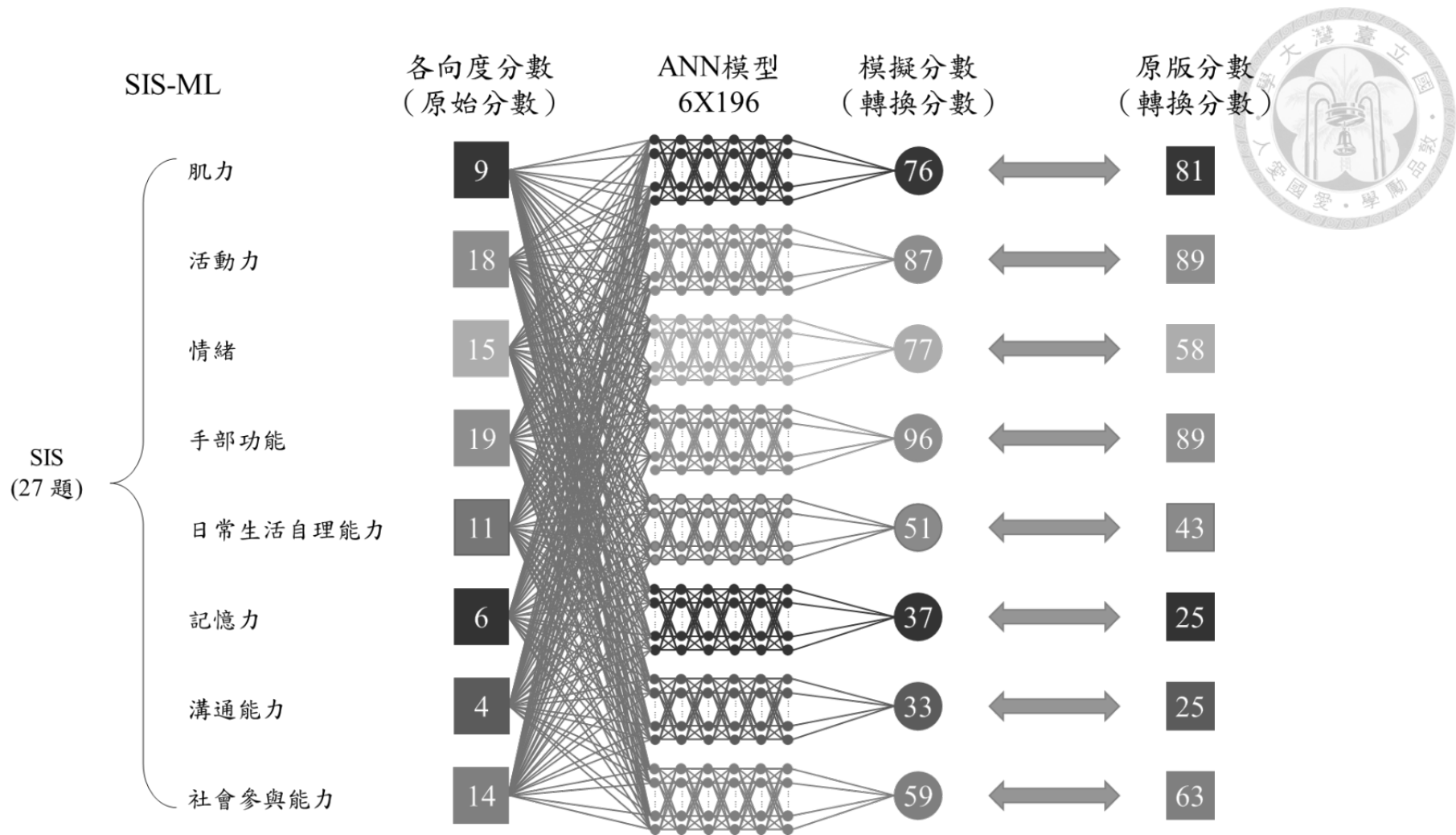
圖八：SIS 各題目對各向度分數之預測程度（模型參數絕對值加總數值越大，預測力越高）



圖九：各 ANN 模型架構於各題組之預測力



圖十：27 題題組之 6X196 模型於測試資料 (n=77) 之模擬分數及原始分數之散佈圖



圖十一：SIS-ML 之 ANN 模型模擬分數示意圖（以單一個案分數為例）



表一：SIS 用於中風患者之信度及反應性彙整

相關研究/心理計量特性	內在一致性 Cronbach's α	再測信度 ICC/ Pearson r /Spearman's ρ	隨機測量誤差 SRD%	反應性 SRM
Chou et al., 2015	(n=263) $\alpha = 0.80 \sim 0.98$	(n=121) ICC= 0.67~0.97	(n=121) SRD%= 12%~68%	
Lin et al., 2010				(n=74) SRM = 0.04~ 0.52
Carod-Artal et al., 2008	(n=174) $\alpha = 0.81 \sim 0.95$	(n=174) ICC= 0.48~0.94	(n=174) SRD%= 30%~81%	
Choi et al., 2017	(n=70) $\alpha = 0.71 \sim 0.96$	(n=70) $\rho = 0.44 \sim 0.79$		

* SIS: Stroke Impact Scale; ICC: Intra-class correlation coefficients; SRD: smallest real difference; SRM: standardized response mean.



表二：SSQOL 用於中風患者之信度及反應性彙整

相關研究/心理計量特性	內在一致性 Cronbach's α	再測信度 ICC/ Pearson r /Spearman's ρ	隨機測量誤差 SRD%	反應性 SRM
Chou et al., 2015	(n=263) $\alpha = 0.65 \sim 0.94$	(n=121) ICC = 0.70~0.91	(n=121) SRD% = 19%~84%	
Lin et al., 2010				(n=74) SRM = 0.01~ 0.17
Muus et al., 2007	(n=152) $\alpha = 0.81 \sim 0.94$	(n=24) r = 0.36~ 0.96		
Pedersen et al.,2018	(n=125) $\alpha = 0.79 \sim 0.97$	(n=36) $\rho = 0.35 \sim 0.94$		

* SSQoL: Stroke-Specific Quality of Life Scale; ICC: Intra-class correlation coefficients; SRD: smallest real difference; SRM: standardized response mean.



表三：以 NIHSS 及 BI 驗證 SIS 各向度收斂效度之數據彙整

相關研究/ Pearson r/ Spearman's ρ /向度	量表	樣本	肌力	活動力	情緒	日常生活 自理能力	手部功能	溝通能力	記憶力	社會參與 能力
Chou et al., 2015	NIHSS	n=263	-0.51	-0.57	-0.30	-0.57	-0.59	-0.26	-0.17	-0.49
Carod-Artal et al., 2008	NIHSS	n=174	-0.56	-0.54	-0.05	-0.64	-0.6	-0.15	-0.09	-0.38
	BI	n=174	0.53	0.79	0.04	0.85	0.55	0.10	0.07	0.47
Choi et al., 2017	NIHSS	n=70	-0.44	-0.47	-0.08	-0.44	-0.51	-0.25	-0.20	-0.52
Vellone et al., 2015	NIHSS	n=45	-0.72	-0.65	-0.27	-0.67	-0.57	-0.25	0.16	-0.50

* BI: Barthel Index; NIHSS: National Institutes of Health Stroke Scale.

表四：6X8 及 6X32 訓練資料 (n=89) 模型分數

模型	6X8																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	0.52	0.39	0.47	0.34	0.33	0.46	0.97	0.97	0.97	0.31	0.96	0.97	0.96	0.96	0.33	0.50	0.40
2	-1.74	-0.88	-1.90	-1.79	-0.32	-1.89	-0.25	-0.87	-0.69	-0.62	-1.80	-1.39	-1.95	-1.24	-1.00	-1.51	-2.06
3	-0.11	0.16	-0.17	-0.12	0.10	-0.30	0.14	0.04	0.02	-0.16	-0.32	-0.20	-0.33	-0.61	-0.27	-0.10	-0.13
4	-4.02	-2.53	-4.27	-3.87	0.95	-3.67	-1.39	-2.53	-2.06	-2.29	-3.95	-3.70	-4.51	-3.62	-3.40	-3.47	-4.76
5	0.69	0.62	0.59	0.64	0.59	0.61	0.53	0.57	0.52	0.48	0.73	0.62	0.78	0.52	0.60	0.66	0.66
6	0.54	0.43	0.52	0.53	0.50	0.49	0.37	0.38	0.35	0.32	0.48	0.51	0.61	0.59	0.49	0.55	0.57
7	1.00	1.00	0.48	0.41	0.98	0.44	0.98	0.99	0.99	0.27	0.99	0.98	0.99	0.98	0.40	0.55	0.54
8	0.50	0.41	0.43	0.49	0.46	0.62	0.40	0.46	0.39	0.29	0.51	0.50	0.48	0.13	0.24	0.45	0.44
模型	6X32																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.99	0.99	0.99	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.96	0.96
2	-0.64	-0.69	-0.44	-0.67	0.94	0.94	-0.66	-0.68	-0.34	-0.87	-0.19	0.94	-0.52	-0.59	-0.65	-0.20	-0.74
3	0.13	0.21	0.28	0.11	0.38	0.97	0.38	0.35	0.41	0.17	0.27	0.37	0.37	0.11	0.18	0.20	0.07
4	-2.48	-2.54	-2.04	-2.38	0.95	0.95	0.95	-2.37	-1.69	-2.71	-1.29	0.92	-2.24	-2.02	-1.78	-1.23	-1.84
5	0.73	0.72	0.65	0.70	0.73	0.79	0.79	0.69	0.68	0.75	0.63	0.73	0.72	0.73	0.70	0.60	0.69
6	0.61	0.57	0.54	0.61	0.99	0.64	0.61	0.57	0.51	0.62	0.50	1.00	0.52	0.61	0.53	0.58	0.61
7	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.98	0.98	0.97
8	0.98	0.48	0.43	0.98	0.98	0.55	0.98	0.57	0.97	0.48	0.49	0.98	0.97	0.49	0.52	0.98	0.97

表五：6X196 及 6X512 訓練資料 (n=89) 模型分數

模型	6X196																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.99	0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
2	0.95	0.95	0.95	0.95	0.94	0.95	0.94	0.94	0.94	0.94	0.94	0.94	0.93	0.93	0.92	0.92	0.92
3	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.96	0.96	0.96	0.97	0.97	0.96	0.97	0.97	0.96	0.97
4	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.94	0.94	0.91	0.92	0.92	0.92	0.90	0.91	0.91
5	0.99	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
6	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.99
7	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98
模型	6X512																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.99	0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.96
2	0.94	0.95	0.94	0.95	0.91	0.92	0.95	0.93	0.94	0.94	0.94	0.93	0.94	0.92	0.89	0.91	0.90
3	0.96	0.96	0.96	0.96	0.96	0.96	0.95	0.95	0.96	0.95	0.97	0.97	0.96	0.97	0.96	0.96	0.96
4	0.95	0.92	0.95	0.93	0.88	0.92	0.93	0.90	0.93	0.92	0.88	0.90	0.89	0.89	0.89	0.89	0.87
5	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
6	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.98	0.98	0.98
8	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98

表六：10X8 及 10X32 訓練資料 (n=89) 模型分數

模型	10X8																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	0.54	0.99	0.99	0.43	0.40	0.43	0.97	0.97	0.44	0.36	0.48	0.97	0.97	0.97	0.32	0.42	0.43
2	-2.19	-0.91	-0.36	-1.17	-1.48	-0.59	-0.29	-0.80	-1.37	-0.08	-0.55	-0.90	-0.89	-0.32	-1.64	-1.22	-1.55
3	-0.27	-0.03	0.19	0.12	-0.17	0.25	0.07	-0.05	0.01	0.13	-0.08	0.19	0.06	-0.16	-0.41	-0.13	-0.05
4	-4.56	-1.90	-2.04	-2.86	-2.79	-2.03	-1.30	-2.74	-3.50	-1.50	-2.87	-2.84	-2.76	-1.91	-3.45	-3.25	-3.86
5	0.53	0.61	0.61	0.47	0.61	0.70	0.49	0.59	0.67	0.57	0.60	0.84	0.71	0.54	0.56	0.60	0.62
6	0.47	0.49	0.52	0.31	0.48	0.50	0.34	0.34	0.48	0.48	0.37	0.55	0.56	0.56	0.41	0.51	0.57
7	0.55	0.47	0.99	0.37	0.98	0.41	0.98	0.35	0.41	0.31	0.99	0.98	0.99	0.98	0.39	0.46	0.46
8	0.43	0.35	0.45	0.48	0.48	0.60	0.42	0.43	0.50	0.42	0.47	0.57	0.51	0.21	0.22	0.41	0.44
模型	10X32																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.99	0.99	0.99	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
2	0.96	0.96	0.95	0.95	0.95	0.95	0.94	0.94	0.94	0.94	0.94	-0.60	0.93	0.91	0.92	0.92	0.93
3	0.97	0.97	0.97	0.96	0.98	0.97	0.96	0.95	0.41	0.23	0.97	0.96	0.94	0.18	0.98	0.97	0.97
4	0.96	0.96	0.95	-2.25	0.95	-2.61	0.95	0.96	0.94	0.93	0.93	0.94	0.90	0.92	0.88	-0.93	0.91
5	0.71	0.67	1.00	1.00	0.72	0.70	0.71	0.72	0.65	0.68	0.72	0.76	0.99	0.73	1.00	0.57	0.66
6	0.60	0.55	1.00	0.61	0.59	1.00	1.00	0.60	0.49	0.56	1.00	0.54	0.99	0.52	0.99	0.99	0.60
7	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	0.98	0.97	0.98

表七：10X196 及 10X512 訓練資料 (n=89) 模型分數

模型	10X196																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.99	0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
2	0.94	0.95	0.94	0.95	0.95	0.95	0.92	0.94	0.93	0.94	0.94	0.94	0.93	0.93	0.93	0.93	0.95
3	0.97	0.97	0.97	0.97	0.98	0.98	0.96	0.97	0.97	0.96	0.97	0.98	0.97	0.97	0.98	0.98	0.98
4	0.93	0.95	0.95	0.95	0.92	0.95	0.95	0.95	0.93	0.95	0.93	0.93	0.93	0.93	0.91	0.92	0.93
5	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.99
6	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.99	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99	0.99
7	0.99	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8	0.99	0.98	0.98	0.99	0.98	0.98	0.98	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
模型	10X512																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	1.00	0.98	0.99	0.98	0.97	0.97	0.97	0.97	0.96	0.97	0.96	0.97	0.97	0.97	0.97	0.97	0.96
2	0.93	< -10	0.88	< -10	0.94	< -10	0.93	0.92	< -10	0.89	< -10	< -10	0.88	0.91	0.83	0.88	0.88
3	0.97	< -10	0.96	< -10	< -10	< -10	0.97	0.96	< -10	0.95	< -10	< -10	0.95	0.96	0.96	0.97	0.97
4	0.92	< -10	0.91	< -10	< -10	< -10	0.93	0.93	< -10	0.86	< -10	< -10	0.86	0.87	0.82	0.86	0.81
5	0.99	0.99	0.98	< -10	< -10	-6.98	0.99	0.99	-6.22	0.98	0.99	0.99	0.99	0.98	0.99	1.00	0.99
6	0.99	-6.70	0.98	-8.66	-8.09	-8.34	0.99	0.99	-8.14	0.99	-6.40	0.99	0.99	0.99	0.99	0.99	0.99
7	1.00	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
8	0.99	0.99	0.98	0.98	0.98	0.99	0.97	0.98	0.99	0.99	0.98	0.99	0.99	0.98	0.98	0.98	0.98

表八：6X196、10X196 及 6X512 驗證資料 (n=90) 之模型分數

模型	6X196																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	0.99	0.95	0.95	0.95	0.85	0.85	0.85	0.84	0.84	0.84	0.84	0.85	0.84	0.84	0.85	0.85	0.85
2	0.91	0.92	0.92	0.92	0.91	0.92	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.87	0.88	0.88	0.77
3	0.82	0.82	0.82	0.81	0.82	0.82	0.77	0.77	0.79	0.79	0.67	0.66	0.68	0.69	0.64	0.60	0.63
4	0.85	0.85	0.84	0.85	0.84	0.85	0.86	0.82	0.82	0.82	0.82	0.74	0.75	0.74	0.74	0.77	0.75
5	0.91	0.91	0.91	0.92	0.92	0.91	0.91	0.90	0.90	0.91	0.92	0.91	0.92	0.92	0.89	0.77	0.75
6	0.96	0.97	0.96	0.96	0.96	0.94	0.95	0.95	0.90	0.89	0.90	0.90	0.89	0.87	0.88	0.88	0.87
7	0.99	0.99	0.99	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
8	0.90	0.90	0.90	0.90	0.90	0.91	0.90	0.90	0.90	0.83	0.81	0.82	0.65	0.68	0.68	0.67	0.74
模型	10X196																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	0.99	0.95	0.95	0.95	0.84	0.85	0.84	0.84	0.84	0.85	0.84	0.85	0.84	0.85	0.85	0.85	0.85
2	0.92	0.92	0.91	0.92	0.91	0.91	0.92	0.92	0.91	0.92	0.93	0.94	0.93	0.88	0.87	0.87	0.77
3	0.81	0.82	0.82	0.80	0.82	0.82	0.76	0.77	0.79	0.78	0.63	0.66	0.68	0.69	0.69	0.64	0.66
4	0.83	0.85	0.84	0.85	0.82	0.84	0.86	0.82	0.80	0.82	0.81	0.72	0.75	0.74	0.73	0.75	0.74
5	0.91	0.91	0.90	0.91	0.92	0.91	0.91	0.91	0.90	0.89	0.91	0.90	0.92	0.92	0.88	0.77	0.76
6	0.96	0.96	0.96	0.96	0.96	0.95	0.94	0.95	0.90	0.88	0.89	0.90	0.90	0.87	0.89	0.89	0.88
7	0.99	0.99	0.98	0.97	0.96	0.96	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
8	0.90	0.90	0.89	0.89	0.90	0.89	0.89	0.90	0.89	0.83	0.82	0.83	0.65	0.68	0.68	0.68	0.73
模型	6X512																
題數	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
1	0.99	0.95	0.95	0.95	0.84	0.85	0.85	0.84	0.84	0.85	0.84	0.85	0.84	0.84	0.84	0.85	0.85
2	0.92	0.92	0.91	0.92	0.91	0.90	0.92	0.91	0.93	0.93	0.93	0.93	0.93	0.87	0.85	0.87	0.77
3	0.81	0.83	0.81	0.82	0.82	0.83	0.77	0.77	0.78	0.78	0.63	0.66	0.70	0.68	0.68	0.66	0.63
4	0.85	0.82	0.84	0.83	0.80	0.83	0.84	0.79	0.82	0.82	0.81	0.73	0.73	0.74	0.74	0.77	0.74
5	0.92	0.91	0.90	0.91	0.90	0.91	0.91	0.90	0.90	0.91	0.92	0.92	0.93	0.93	0.89	0.79	0.77
6	0.97	0.96	0.96	0.96	0.97	0.95	0.95	0.94	0.91	0.90	0.90	0.90	0.90	0.89	0.89	0.89	0.89
7	0.99	0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
8	0.90	0.89	0.90	0.91	0.89	0.91	0.91	0.90	0.91	0.82	0.82	0.81	0.65	0.68	0.65	0.69	0.71

表九：6X196、10X196 及 6X512 之訓練 (n=89)、驗證 (n=90)、測試 (n=77) 資料之模型分數

模型	6X196					
題數	32	31	30	29	28	27
組別	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試
1	1.00 0.99 0.99	0.99 0.95 0.97	0.99 0.95 0.97	0.98 0.95 0.97	0.97 0.85 0.89	0.97 0.85 0.89
2	0.95 0.91 0.90	0.95 0.92 0.91	0.95 0.92 0.91	0.95 0.92 0.91	0.94 0.91 0.90	0.95 0.92 0.91
3	0.97 0.82 0.85	0.97 0.82 0.86	0.97 0.82 0.85	0.97 0.81 0.85	0.97 0.82 0.84	0.97 0.82 0.83
4	0.95 0.85 0.87	0.95 0.85 0.87	0.95 0.84 0.87	0.95 0.85 0.85	0.95 0.84 0.86	0.95 0.85 0.85
5	0.99 0.91 0.88	0.99 0.91 0.89	0.99 0.91 0.89	1.00 0.92 0.89	0.99 0.92 0.89	0.99 0.91 0.89
6	0.99 0.96 0.96	0.99 0.97 0.96	0.99 0.96 0.96	0.99 0.96 0.96	0.99 0.96 0.96	0.99 0.94 0.95
7	1.00 0.99 0.99	1.00 0.99 0.99	0.99 0.99 0.99	0.99 0.97 0.97	0.99 0.97 0.96	0.99 0.97 0.96
8	0.98 0.90 0.83	0.98 0.90 0.85	0.98 0.90 0.85	0.98 0.90 0.84	0.98 0.90 0.84	0.98 0.91 0.85
模型	10X196					
題數	32	31	30	29	28	27
組別	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試
1	1.00 0.99 0.99	0.99 0.95 0.97	0.99 0.95 0.97	0.98 0.95 0.97	0.97 0.84 0.89	0.97 0.85 0.89
2	0.94 0.92 0.90	0.95 0.92 0.90	0.94 0.91 0.90	0.95 0.92 0.90	0.95 0.91 0.90	0.95 0.91 0.90
3	0.97 0.81 0.85	0.97 0.82 0.85	0.97 0.82 0.85	0.97 0.80 0.84	0.98 0.82 0.83	0.98 0.82 0.84
4	0.93 0.83 0.86	0.95 0.85 0.86	0.95 0.84 0.85	0.95 0.85 0.86	0.92 0.82 0.85	0.95 0.84 0.84
5	1.00 0.91 0.88	1.00 0.91 0.89	1.00 0.90 0.88	1.00 0.91 0.90	1.00 0.92 0.90	1.00 0.91 0.90
6	0.99 0.96 0.95	0.99 0.96 0.95	0.99 0.96 0.95	0.99 0.96 0.96	0.99 0.96 0.96	1.00 0.95 0.95
7	0.99 0.99 0.99	1.00 0.99 0.99	0.99 0.98 0.99	0.99 0.97 0.96	0.99 0.96 0.96	0.99 0.96 0.96
8	0.99 0.90 0.84	0.98 0.90 0.85	0.98 0.89 0.82	0.99 0.89 0.83	0.98 0.90 0.85	0.98 0.89 0.83
模型	6X512					
題數	32	31	30	29	28	27
組別	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試	訓練 驗證 測試
1	1.00 0.99 0.99	0.99 0.95 0.97	0.99 0.95 0.97	0.98 0.95 0.97	0.97 0.84 0.89	0.97 0.85 0.89
2	0.94 0.92 0.91	0.95 0.92 0.92	0.94 0.91 0.91	0.95 0.92 0.92	0.91 0.91 0.90	0.92 0.90 0.90
3	0.96 0.81 0.85	0.96 0.83 0.85	0.96 0.81 0.85	0.96 0.82 0.85	0.96 0.82 0.85	0.96 0.83 0.84
4	0.95 0.85 0.86	0.92 0.82 0.86	0.95 0.84 0.86	0.93 0.83 0.84	0.88 0.80 0.85	0.92 0.83 0.86
5	0.99 0.92 0.89	0.99 0.91 0.89	0.99 0.90 0.88	0.99 0.91 0.89	0.99 0.90 0.88	0.99 0.91 0.89
6	0.99 0.97 0.96	0.99 0.96 0.96	0.99 0.96 0.96	0.99 0.96 0.95	0.99 0.97 0.96	0.99 0.95 0.95
7	0.99 0.99 0.99	0.99 0.99 0.99	0.99 0.98 0.99	0.99 0.97 0.97	0.99 0.97 0.97	0.99 0.97 0.97
8	0.98 0.90 0.83	0.98 0.89 0.83	0.98 0.90 0.83	0.98 0.91 0.86	0.98 0.89 0.84	0.98 0.91 0.85

表十：27 至 32 題題組之訓練 (n=89)、驗證 (n=90)、測試 (n=77) 資料之 8 個向度
平均模型分數

模型/題組	32		31		30		29		28		27	
	訓練	0.98	訓練	0.98	訓練	0.98	訓練	0.98	訓練	0.97	訓練	0.98
6X196	驗證	0.92	驗證	0.91	驗證	0.91	驗證	0.91	驗證	0.90	驗證	0.90
	測試	0.91	測試	0.91	測試	0.91	測試	0.90	測試	0.90	測試	0.89
	訓練	0.98	訓練	0.98	訓練	0.98	訓練	0.98	訓練	0.97	訓練	0.98
10X196	驗證	0.91	驗證	0.91	驗證	0.91	驗證	0.91	驗證	0.89	驗證	0.89
	測試	0.91	測試	0.91	測試	0.90	測試	0.91	測試	0.89	測試	0.89
	訓練	0.98	訓練	0.97	訓練	0.97	訓練	0.97	訓練	0.96	訓練	0.97
6X512	驗證	0.92	驗證	0.91	驗證	0.91	驗證	0.91	驗證	0.89	驗證	0.89
	測試	0.91	測試	0.91	測試	0.91	測試	0.91	測試	0.89	測試	0.89
	訓練	0.98	訓練	0.97	訓練	0.97	訓練	0.97	訓練	0.96	訓練	0.97



表十一：SIS-ML 於測試資料 (n=77) 之平均值、標準誤及隨機測量誤差

向度	SIS		SIS-ML		SIS 與 SIS-ML 之差異絕對值		差異百分比
	平均值	標準差	平均值	標準差	平均數	標準差	
肌力	47.0	30.4	48.7	27.9	7.6	6.5	16%
記憶力	85.2	20.5	86.7	16.2	3.9	4.9	5%
情緒	69.0	20.6	70.5	18.3	6.6	5.1	10%
溝通能力	90.2	17.3	92.0	14.5	4.0	5.2	4%
日常生活自理	69.2	27.4	71.1	24.8	6.3	6.3	9%
活動力	65.4	31.1	67.4	29.2	5.1	4.4	8%
手部功能	51.4	41.5	52.5	38.9	5.5	5.5	11%
社會參與能力	64.8	27.8	62.3	26.7	7.9	7.5	12%

* SIS: Stroke Impact Scale

表十二：SIS-ML 於測試資料之同時效度及收斂效度

Domain	原版 SIS	NIHSS	BI
肌力	0.95	-0.54	0.43
記憶力	0.97	-0.34	0.23
情緒	0.92	-0.42	0.25
溝通能力	0.94	-0.35	0.22
日常生活自理	0.95	-0.59	0.56
活動力	0.98	-0.57	0.74
手部功能	0.98	-0.56	0.41
社會參與能力	0.92	-0.58	0.44

* SIS: Stroke Impact Scale; BI: Barthel Index; NIHSS: National Institutes of Health Stroke Scale.

附錄一：SIS-ML 各向度之候選題目（依預測力順序）

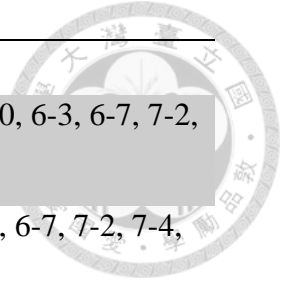
向度/題目代號*	預測力順序			
	1	2	3	4
肌力	1-3	1-4	1-1	1-2
記憶力	2-2	2-1	2-7	2-6
情緒	3-8	3-3	3-5	3-1
溝通能力	4-4	4-1	4-2	4-5
日常生活自理	5-2	5-5	5-10	5-8
活動力	6-3	6-7	6-5	6-2
手部功能	7-2	7-4	7-3	7-1
社會參與能力	8-3	8-6	8-1	8-8

*題目代號為[向度-題號]組成。向度順序依原版 SIS 之向度順序依序編號。題號為原版 SIS 之題號，每向度皆自 1 開始編號。各向度之題數：肌力 4 題，記憶力 7 題，情緒 9 題，溝通能力 7 題，日常生活自理 10 題，活動力 9 題，手部功能 5 題及社會參與能力 8 題。

附錄二： 17 組 SIS-ML 候選題組之題目



題組	項目 (以題目代號呈現)
32 題	1-3, 1-4, 1-1, 1-2, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 6-2, 7-2, 7-4, 7-3, 7-1, 8-3, 8-6, 8-1, 8-8
31 題	1-3, 1-4, 1-1, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 6-2, 7-2, 7-4, 7-3, 7-1, 8-3, 8-6, 8-1, 8-8
30 題	1-3, 1-4, 1-1, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 6-2, 7-2, 7-4, 7-3, 8-3, 8-6, 8-1, 8-8
29 題	1-3, 1-4, 1-1, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 6-2, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
28 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 6-2, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
27 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 3-1, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
26 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 4-4, 4-1, 4-2, 4-5, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
25 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 4-4, 4-1, 4-2, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 6-5, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
24 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 4-4, 4-1, 4-2, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 7-2, 7-4, 8-3, 8-6, 8-1, 8-8
23 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 3-5, 4-4, 4-1, 4-2, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 7-2, 7-4, 8-3, 8-6, 8-1
22 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 4-4, 4-1, 4-2, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 7-2, 7-4, 8-3, 8-6, 8-1
21 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 7-2, 7-4, 8-3, 8-6, 8-1
20 題	1-3, 1-4, 2-2, 2-1, 2-7, 2-6, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7, 7-2, 7-4, 8-3, 8-6
19 題	1-3, 1-4, 2-2, 2-1, 2-7, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 5-10, 5-8, 6-3, 6-7,



7-2, 7-4, 8-3, 8-6

18 題 1-3, 1-4, 2-2, 2-1, 2-7, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 5-10, 6-3, 6-7, 7-2,
7-4, 8-3, 8-6

17 題 1-3, 1-4, 2-2, 2-1, 2-7, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 6-3, 6-7, 7-2, 7-4,
8-3, 8-6

16 題 1-3, 1-4, 2-2, 2-1, 3-8, 3-3, 4-4, 4-1, 5-2, 5-5, 6-3, 6-7, 7-2, 7-4, 8-3,
8-6
