|  |
| --- |
| **國家科學及技術委員會補助專題研究計畫報告** |
|

**以口語表達與手寫表現探討自閉症兒童的行為特徵與學習輔助**

報告類別：成果報告

計畫類別：個別型計畫

計畫編號：MOST 112-2221-E-468-009-

執行期間：112 年 08 月 01 日 至 113 年 12 月 31 日

執行機構及系所：亞洲大學資訊工程學系

計畫主持人：陳良弼

共同主持人：藍先元

|  |
| --- |
| 本研究具有政策應用參考價值： □否 ■是，建議提供機關衛生福利部  (勾選「是」者，請列舉建議可提供施政參考之業務主管機關)  本研究具影響公共利益之重大發現：■否 □是 |

中 華 民 國 112 年 12 月 1 日

中文摘要: 隨著資訊電子化的政策及便利性，醫療單位朝向透過電子病歷記錄病的記錄而 醫療資料經年累月累積形成巨量資料，帶來巨量醫療資料的整合、處理與分析等 研究議題。現今社會忙碌的人也因為網路資訊的發達，可以透過線上諮詢取得專 業醫生的回覆，清楚是否該就醫或獲得其他醫療訊息。而多元的社群媒體平台， 讓人們可以藉此查詢與學習各種知識，也能夠在平台上創作與分享各種內容，造 就社群媒體成為現代生活不可或缺的核心媒介。因此，透過分析醫療資料、社群 媒體資料與問答網站資料，可以產生各式新型醫療服務，不僅能幫助普通民眾了 解自身需求，也能協助醫院與醫生提升病患醫療滿意度，降低醫療資源的浪費。 著眼於此，本計畫以三年為期，探討運用多元資料與深度學習技術於新型醫療服 務之開發。在本年度計畫執行過程中，我們已完成『基於組合特徵的注意力機制 與完全共享的多任務學習之生物醫學專名識別』、『結合知識檢索來進行心理健康 狀況檢測』、『使用字典選擇的高風險文章的詞級模型進行自殺風險評估』、『透過 整合對話歷史和輔導者特徵進行情緒輔導對話的策略預測』共四項研究。本期末 報告茲就本年度所完成的研究成果進行報告。

中文關鍵詞: 輔助醫療決策、資料分析、問答模型、疾病預測模型、情感分析

英文摘要: The convenience of the Internet has made most medical institutions to build electronic health records (EHR). The accumulated EHR over time provides good opportunities for new research. Moreover, people nowadays can easily get online professional medical consultations via Internet in addition to traditional medical treatments. Finally, social media have become an indispensable core media for most people. In this project, we research into medical service applications by analyzing EHR, social media, and Q&A data. The expected results not only assist people on what they need, but also help medical institutions and medical professionals improve the satisfaction of patients and reduce the waste of medical resources. In this progress report, four research results we achieved in this year are presented, including 1) biomedical named entity recognition with the combined feature attention and fully-shared multi-task learning, 2) the detection of mental health conditions by incorporating knowledge retrieval, 3) suicide risk assessment using word-level model with dictionary-based risky posts selection, and 4) predicting the following support strategy during the emotional support dialogue by integrating the dialogue history and supporter features.

英文關鍵詞: Medical Decision Support, Data Analysis, QA Model, Disease Prediction Model, Sentiment Analysis

**科技部補助專題研究計畫成果報告**

**（□期中進度報告 / ■期末報告）**

**以口語表達與手寫表現探討自閉症兒童的行為特徵與學習輔助**

**Identification of Behavioral Characteristics of Autism Children with Their Narrative and Handwriting**

計畫類別：■ 個別型計畫　　□ 整合型計畫

計畫編號：112-2221-E-468-009-

執行期間：2023年08月01日至2024年12月31日

執行機構及系所：亞洲大學資訊工程學系

計畫主持人：陳良弼

計畫參與人員：蔡昀陞、林鎰鋒、林昀昇、嚴翎愷、李昀叡、SYAUKI AULIA THAMRIN

本計畫除繳交成果報告外，另含下列出國報告，共 \_0\_ 份：

□ 執行國際合作與移地研究心得報告

□ 出席國際學術會議心得報告

□ 出國參訪及考察心得報告

中 華 民 國　　112 　年　　12　　月 　 1　　 日

**中文摘要**

隨著資訊電子化的政策及便利性，醫療單位朝向透過電子病歷記錄病人的看診記錄，而醫療資料經年累月累積形成巨量資料，帶來巨量醫療資料的整合、處理與分析等研究議題。現今社會忙碌的人也因為網路資訊的發達，可以透過線上諮詢，取得專業醫生的回覆，清楚是否該就醫或獲得其他醫療訊息。而多元的社群媒體平台，讓人們可以藉此查詢與學習各種知識，也能夠在平台上創作與分享各種內容，造就社群媒體成為現代生活不可或缺的核心媒介。因此，透過分析醫療資料、社群媒體資料與問答網站資料，可以產生各式新型醫療服務，不僅能幫助普通民眾了解自身需求，也能協助醫院與醫生提升病患醫療滿意度，降低醫療資源的浪費。著眼於此，本計畫以三年為期，探討運用多元資料與深度學習技術於新型醫療服務之開發。在本年度計畫執行過程中，我們已完成『基於組合特徵的注意力機制與完全共享的多任務學習之生物醫學專名識別』、『結合知識檢索來進行心理健康狀況檢測』、『使用字典選擇的高風險文章的詞級模型進行自殺風險評估』、『透過整合對話歷史和輔導者特徵進行情緒輔導對話的策略預測』共四項研究。本期末報告茲就本年度所完成的研究成果進行報告。

**Abstract**

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social communication and interaction, as well as restricted and repetitive behaviors. ASD affects individuals across various aspects of their lives, including their cognitive, social, and motor skills. Early detection and intervention are crucial for improving outcomes and providing appropriate support to individuals with ASD. With recent advancement regarding computational language models, we aim to utilize machine learning and deep neural network models to analyze children's narratives and handwriting to detect autism tendencies and characteristics. These approaches can contribute to improving diagnostic efficiency and providing comprehensive language evaluations to better understand and support the needs of children with autism. In this progress report, three research results we achieved in this year are presented, including 1) identifying Chinese handwriting characteristics for detecting children with autism, 2) analyzing handwriting characteristics of children with autism by Chinese characters and Mandarin Phonetic Symbols I, 3) using natural language processing to analyze autistic tendencies from children's picture book narratives, with the second research being the successor of the first.

**一、前言**

自閉症譜系障礙（Autism Spectrum Disorder, ASD）是一種複雜的神經發育障礙，主要特徵為在社交互動和溝通方面的挑戰，並伴隨著限制性、重複性的行為模式、興趣或活動 [2-1]。這是一種終身的障礙，通常從幼年就開始影響個體在各個功能領域的表現。「譜系」一詞指的是自閉症譜系障礙中症狀、能力和特徵的廣泛變化，這些特徵在不同ASD患者之間差異很大。

根據自閉症和發展障礙監測（Autism and Developmental Disabilities Monitoring, ADDM）報告中顯示，2020年美國8歲兒童中每千人有27.6例被診斷為ASD，大約每36名兒童中就有一名被診斷為ASD。並且此研究顯示ASD患者性別差異顯著，男性的ASD發生率是女性的約3.8倍（Maenner等，2023）。2023年，台灣最新的自閉症人數為19,078人，其中74.6%被認定為輕度ASD。大多數輕度ASD兒童通常在進入小學後才被發現並診斷。這些自閉症兒童可能會面臨各種學習和社交挑戰，並需要個別的教育支持。因此，幫助學校教師發現具有自閉症特徵的兒童，可能會加快提供適當教育資源的進程。

診斷ASD的過程漫長且耗時。其中一個原因是因為ASD的定義主要依賴行為，而非基因上的特徵或物理生物標記（Schaaf等，2020）。ASD的診斷過程通常涉及多方面的評估。其主要分為兩個階段：篩查和綜合診斷評估。典型的ASD診斷過程包括多學科的合作，並收集來自各種來源的信息，包括但不限於：(1) 與父母或照護者的訪談，(2) 幼兒園或學校教師的教育日誌或行為觀察，(3) 標準化的行為評估和體格檢查（Goldstein & Ozonoff, 2018）。整個診斷過程通常需要數月的時間，並伴隨著多次的醫院訪問，這對於許多家庭而言可能是巨大的時間和資源負擔。

自閉症個體的一個顯著特徵是他們在不同的社交情境中，持續面臨社交溝通和互動的挑戰（美國精神醫學協會，2022）。在不同語言中，研究一再表明，自閉症個體在理解和社交性地使用語言方面存在缺陷（Engberg-Pedersen & Christensen, 2016；Sah & Torng, 2015；Mäkinen等，2014）。自閉症兒童講故事時的敘事能力是在語言使用方面最常見的問題（Baixauli, Colomer, Roselló, & Miranda, 2016）。研究發現自閉症兒童的敘事往往缺乏連貫性和因果聯繫，並可能包含無關或不符合的內容（Diehl, Bennetto, & Young, 2006）。此外，與同齡典型發展的兒童相比，自閉症兒童在講故事時的語句數量較少，詞彙多樣性也較低（Capps, Losh, & Thurber, 2000）。自閉症兒童無法描述角色的思想和情感，其原因可能是他們無法理解角色行為背後的動機。綜合以上研究與ASD兒童的敘事特徵，我們可能可以通過兒童的敘事來識別具有自閉症特徵的兒童。

另一方面，自閉症兒童在將連續的動作轉化為整合性動作時可能也會遇到困難。這些精細運動技能的困難會影響他們的日常生活，導致挫折。手眼協調的缺陷可能影響需要精確度的活動，包括書寫。

[研究] 本項目旨在發現手寫和敘事中的潛在特徵，分別反映出前述的兩項發現。我們已經完成了名為「[名稱]」的研究。方法和結論將在後續部分中介紹。

**國內外相關研究**

1. **ADOS-2**

**2、Deep Learning for the Automatic Detection of ASD**

**3、Handwriting Characteristics of ASD Children**

**4、****Computational models for analyzing narration**

**二、研究方法、進行步驟及執行進度報告**

我們擬探討多元資料與深度學習技術於新型醫療服務之開發，並著重於三大主軸：『醫療資料』、『社群媒體』、『問答網站』。第三年成果報告如下所示:

**醫療資料之成果報告：基於組合特徵的注意力機制與完全共享的多任務學習之生物醫學專名識別**

**研究目的**

生物醫學研究的快速發展導致生物醫學文獻數量呈爆炸性指數增長，使得生物醫學學者難以跟上尖端研究的步伐。因此，迫切需要有效的自然語言處理（NLP）工具來檢索、組織和管理龐大的生物醫學數據和信息。NLP工具在文本挖掘中充當機器和人類語言之間的重要橋樑。NLP系統在不同目標上有各種應用，其中命名實體識別（NER）是最基本和重要的任務之一。

生物醫學命名實體識別（BioNER）是任何生物醫學文獻挖掘任務的首要步驟，其目標是檢測生物醫學實體的邊界並預測其實體類型，如疾病、基因、物種、化學物等。BioNER系統的性能直接影響到下游應用，如生物醫學關係提取、藥物相互作用任務和知識庫構建等。

近年來，深度神經網路，尤其是預訓練的語言模型，在生物醫學專名識別領域取得了令人巨大的進展。然而，由於缺乏大規模高品質的注釋資料和領域知識，其性能仍然有限。為了解決這個問題，我們提出了一個新的基於預訓練的BioBERT的多任務學習模型；該模型帶有一個新的注意力模組，可以將自動處理過的句法資訊集成到模型中。

**研究方法**

我們單任務模型的整體架構詳細描述於圖1。左側描述了用於BioNER序列標記範例的架構主幹，右側則是處理語法信息的過程。我們提出了一個新型的注意模塊，將語法信息整合到模型的主體部分。

為了獲取語法信息，首先，我們運行開源的自然語言處理工具包，例如Stanford CoreNLP Toolkit（SCT）[38]，以獲取輸入一張含有 文字, 圖表, 螢幕擷取畫面, 字型 的圖片

自動產生的描述句子的分析結果。然後，我們從分析結果中提取每個單詞在中的上下文特徵及其相應的句法標籤。

圖1、單任務模型架構

一張含有 文字, 螢幕擷取畫面 的圖片

自動產生的描述 受到[7]的啟發，我們採用注意力方法將語法特徵融入到BioBERT模型中，一旦從組合特徵注意力模塊獲得輸出向量，我們將其送至全連接層，然後是一個softmax層，對於句子中的每個單詞，標記的機率分布可以表示為以下形式:

圖2、多任務模型架構

，其中[表示每種BioNER標籤類型的概率，例如，”B”、”I”和”O”，其中”W”和”b”是可訓練的參數。

多任務學習（MTL）成功應用於解決BioNER任務中有限標註數據的挑戰。然而，先前的MTL模型通常同時使用多個數據集訓練龐大模型，其中一些參數共享，而其他則分離且特定於任務。這導致模型參數急遽擴大，且在不同數據集之間分享信息受限。受到[10]的啟發，我們提出了一種簡單而有效的MTL方法，如圖2所示，每個輸入詞序列的開始和結束處都附有一對標籤標識符：“<tag>”和“</tag>”，其中“tag”表示輸入句子所屬的數據集名稱。在訓練中，將不同數據集的句子與標籤標識符一起輸入模型。這些標籤標識符區分每個句子的來源，影響每個詞的隱藏表示。由於只在編碼之前更改輸入句子，無需修改整個模型結構，這種全面共享的MTL方法能夠在訓練步驟中共享所有參數，整合不同數據集，並在無需特定於任務的層的情況下訓練模型。此外，我們訓練模型同時處理多種生物醫學實體類型的數據集，以捕捉跨類型的生物醫學信息。標籤標識符“<tag>”使用數據集的名稱，以捕捉由於構一張含有 文字, 螢幕擷取畫面, 數字, 字型 的圖片

自動產生的描述造和標註規則的不同而存在的數據集差異。

表1、單任務模型效能比較

**實驗結果**

我們在七個公開可用且在先前的研究中廣泛使用的BioNER基準數據集上進行實驗。對於每個數據集，我們採用了與Crichton等人[8]相同的訓練、驗證和測試集分割策略。

用於比較的單任務模型有三種：首先是由Lee等人提出的基本BioBERT[4]模型，其次是基於kvMN的BioKMNER模型[5]。此外，我們實現了一個新的注意機制，即雙向注意（TWA），集成到BioKMNER模型中，形成“BioBERT-TWA”模型，預期性能優於BioKMNER。這三個模型使用相同的語法信息類型，並使用相同的NLP工具包SCT獲取語法信息。

表1為我們所提出的模型(BioBERT-CFA)和其他模型的比較，我們能從表中觀察出以下幾點:

首先，和最普通沒有使用語法信息的BioBERT比較，其餘有使用語法信息的模型都獲得了較好的效果，這說明了利用語法信息有助於生物醫學專名識別，此外，我們提出的單任務模型BioBERT-CFA在所有基準數據集上均取得了最佳性能，並通過整合句法信息對基線一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述進行了顯著增強。例如，與BioBERT相比，BioBERT-CFA在JNLPBA、Species-800和NCBI-disease數據集上分別取得了1.06%、0.95%和0.80%的F1分數提升，這證實了所提出的CFA模塊的有效性和普遍性。與BioBERT-TWA相比，BioBERT-CFA模型使用了更多的新型注意力機制和嵌入方法，因此BioBERT-CFA能夠提供卓越的性能。

表2、多任務模型效能比較

我們使用全面共享的多任務學習方法在基本BioBERT模型上對所有數據集進行訓練，形成BioBERT-MTM和BioBERT-CFA-MTM。BioBERT-CFA-MTM使用依賴關係（Dep）作為語法信息。 BioBERT-STM是單一數據集訓練的基準模型。為了比較，我們設計了BioBERT-DM模型，直接在整個混合數據集上訓練BioBERT，結果如表2所示，其性能大幅下降。BioBERT-MTM在所有數據集上相對於基準模型都取得穩定的改善，證實了全面共享的多任務學習方法能夠獲得比單一數據集更好的性能。BioBERT-MTM在BC5CDR-Chemical數據集上僅提高0.12％，這是因為BC5CDR-Chemical是七個數據集中唯一的化學類型。BioBERT-CFA-MTM在所有數據集上都表現出色，進一步驗證了提出的CFA模塊和MTL方法的有效性。

**社群媒體之成果報告：結合知識檢索來進行心理健康狀況檢測**

**研究目的**

心理健康狀況的問題日益嚴重，不但患者本身面臨巨大的風險，同時也為全球帶來極高的經濟負擔[39-42]。一些研究顯示，如果患者可以被及早發現，將可以減少精神疾病對我們社會的影響，並降低經濟成本[43, 44]。所以如何盡早偵測到心理疾病就成為很關鍵的課題。而近年來越來越普及且貼近人們生活的社群媒體，就成為研究者們常用來預測心理健康狀況的資料來源[45-47]。然而，過去的研究多只著重於使用者在社群媒體上發布的內容[14, 15]，但醫生判斷精神疾病還會依靠專業的知識和測驗工具[48, 49]。所以本研究希望能藉由導入這些工具和知識，來有效的提升模型對於心理健康狀況的預測效能。

一張含有 文字, 螢幕擷取畫面 的圖片

自動產生的描述**研究方法**

我們從心理健康的檢測工具、維基百科心理健康相關的條目、以及心理健康的權威書籍DSM-5[49]中收集心理健康相關的外部知識的片段。並在此之上，我們使用大量且不設限的維基百科資料，當成普通常識，來測試加上普通常識能否提升心理健康狀況的預測。我們藉由預訓練語言模型來特徵提取使用者發文和外部知識片段的高維表示，並使用DPR[21]來檢索與使用者發文相關的外部知識。DPR是利用發文內容與知識片段的高維表示間的內積，來計算出兩者之間的相似度，並最終獲得與發文內容相關的知識片段。找到相關的知識片段後，使用幾種經典的注意力機制[50]，來建構我們的深度學習模型如圖3。我們對於資料集中的九種精神健康狀況，都使用二元分類的方式對每種心理健康狀況進行預測，最後可以得到每個用戶是否患有該心理健康狀況的結果。

圖3、模型架構

**實驗結果**

我們與過去的研究[11, 51]的F1分數進行比較，並在所有心理健康狀況的預測上，都超越了過去研究的實驗結果。結果如同表3所示，在表中我們可以發現到，單純使用預訓練語言模型來對發文內容進行特徵提取，並且加上注意力機制，就已經能夠打敗過去研究的模型效能。而引入心理相關的知識片段，則可以更進一步的提升模型的效能。最終在思覺失調症的預測上，比過去最好的結果高出了10%以上，即使扣除預訓練語言模型的影響，也高出將近3%。這個結果顯示，引入心理相關的知識可以有效的提升心理健康狀況的預測，這對於提早偵測心理疾病這個任務，是一個很有意義的結果。

**社群媒體之成果報告：使用字典選擇的高風險文章的詞級模型進行自殺風險評估**

表3、與不同方法的比較

**研究目的**

自殺是個在全球都很嚴重的議題，根據美國CDC指出在美國約有1200萬人有過自殺想法，320萬人有過自殺計畫，120萬人嘗試過自殺。在台灣，2020年則有3656人死於自殺，40432人嘗試過自殺[52]。

隨著資訊科技的蓬勃發展，因為社群媒體有著較好的匿名性，許多人願意把內心不常與別人透漏的想法，透過社群平台紀錄與分享，使得社群論壇的資料被廣泛運用來進行自殺風險相關的研究。以往的研究透過文章級別的注意力機制來關注高風險的文章，但這並非每次都可以奏效，這樣子的問題在使用特徵提取的文章向量中更為嚴重，因為文章在模型訓練階段以前就已經轉換成向量，也已經失去了文字級別的資訊。因此本研究中我們先使用自殺字詞字典來篩選出可能被文章級別的注意力機制所忽略的高風險文章，並且對選出的高風險文章運用文字級別的模型來找回因為特徵提取的文章向量所失去的資訊。

**研究方法**

每個用戶都有帖子，按照它們的發帖時間順序排列，其中表示最近的帖子。帖子中的包含詞語。給定用戶的一系列帖子，我們的目標是構建一個模型，將用戶分類為五個不同水平的自殺風險，即SU（自殺）、IN（指示）、ID（構想）、BR（行為）和AT（嘗試）。

我們模型PLWL的架構如圖4所示。PLWL包含一個基於帖子的模型（PL）和一個基於詞語的模型（WL）。PL使用用戶的所有帖子作為輸入，並在第一階段進行訓練。然後將權重儲存並在第二階段中使用。在第二階段，所有帖子都經過PL以獲取用戶表示a。在WL中，帖子經過一個帖子選擇層，選擇高風險的帖子作為以下層的輸入，以獲取高風險的帖子表示b。然後將這兩個表示a和b傳遞到融合層以獲得最終的預測結果。

一張含有 文字, 圖表, 螢幕擷取畫面, 平行 的圖片

自動產生的描述**實驗結果**

表4、和不同方法的比較結果

圖4、PLWL模型架構

 我們將我們的模型與Contextual CNN[25]、SDM [53]、ContextBERT [24]、SISMO [26]（進行了50次運行和10次運行的比較）以及SASI [27]（進行了10次運行的比較）進行比較。結果如表4所示，無論是50次運行還是10次運行，PLWL與SMKD(suicide method keyword dictionary)字典達到了與先前工作相同的性能。PLWL（SMKD）可以通過增加GR來提高FScore。

表5、帶有/不帶有過採樣的關鍵詞字典比較

表5顯示了具有不同關鍵詞字典的PLWL的性能，其在macro F1分數上至少比SISMO高出0.86%。PLWL (3ST)的性能優於PLWL (SMKD)，因為3ST是一個明確定義的自殺企圖詞典，而SMKD是一個自殺方法詞典。CKD結合了SMKD和3ST的優勢，表現最佳。

由於AT用戶的數量僅占總用戶數的9％，我們對AT用戶進行了10％的過採樣，以減少數據不平衡的程度。請注意，只有AT用戶被過採樣，因為其他標籤相對平衡。如表5所示，過採樣方法使PLWL的性能提高了近1％。總之，具有對AT進行10%過採樣的PLWL（CKD）獲得了0.4203的macro F1分數，這是最佳結果。

**問答網站之成果報告：透過整合對話歷史和輔導者特徵進行情緒輔導對話的策略預測**

**研究目的**

世界上大約 20% 的人口患有心理健康疾病。然而由於心理衛生相關專業人員的普遍短缺，如何有效地獲得心理健康輔導（如諮商或心理治療）仍是一項全球性挑戰 [54]。隨著科技的快速發展，透過聊天機器人提供心理健康輔導的解決方案被提出，希望能夠增加心理健康服務的可及性並減少心理治療師的工作量。

近年來，越來越多的研究聚焦在這類聊天機器人在提供心理健康輔導時應該考慮哪些社交技能或心理治療策略。[36] 提出了Emotional Support Conversation的框架，包含了數種基於Helping Skill Theory的情緒輔導策略。該研究還發現使用情緒輔導策略能夠更好地安慰患者，並表明在正確的時間使用情緒輔導策略對於找出患者的問題至關重要。儘管情緒輔導策略的重要性已經得到了證明，但如何在對話中適當地使用這些策略仍然是一個很大的挑戰。為了解決這個問題，我們的目標是引入深度學習技術來幫助我們決定在對話中該使用何種情緒輔導策略。並期望通過這項任務，協助輔導者根據建議的情緒輔導策略在不同時間提供適當的回應，從而有效地提供心理健康輔導。

**研究方法**

一張含有 文字 的圖片

自動產生的描述在我們的任務中，我們旨在透過深度學習方法來預測情緒輔導對話期間所使用的情緒輔導策略。如圖5所示，我們從情緒輔導對話的對話歷史(Dialogue History)中提取對話片段(Dialogue Fragment)來進行預測。其中對話片段所包含的話語數量由滑動窗口(Sliding Window)的大小決定，並作為研究預測性能的參數來進行調整，並且每個對話片段都會有一個對應的情緒輔導策略。我們根據這些對話片段來預測相應的情緒輔導策略，並將其表述為一個多分類的問題。

圖5、情緒輔導對話

為此我們提出了一種基於深度學習的方法來解決這個問題。除了對話片段所包含的資訊外，我們進一步從情緒輔導對話中提取三個與輔導者相關的特徵，以提高模型預測的準確性。如圖6所示，我們的模型主要由兩個部分所組成：(1) Utterance Encoder，它學習由連續話語所組成的對話片段的表徵。(2) Structure Encoder，它學習情緒輔導對話中與輔導者相關的表徵，其中包含三個特徵：Strategy Feature、Orientation Feature和Position Feature。我們將兩者學習到的表徵合併，並用於預測之後對話應該使用的情緒輔導策略。

一張含有 文字, 螢幕擷取畫面, 軟體, 多媒體軟體 的圖片

自動產生的描述**實驗結果**

表6、不同滑動窗口大小對模型性能的影響

圖6、模型架構

對話片段是從情緒輔導對話的對話歷史中提取的，其中包含的話語數量受滑動窗口大小的限制。在我們的實驗中，測試了滑動窗口大小從1到32的對話片段，以及整個對話歷史。為此我們將包含整個對話歷史的對話片段的滑動窗口大小設置為100，因為資料集中最長的情緒輔導對話是由99個話語所組成。如表6所示，我們發

現模型的性能起初隨著滑動窗口大小的增加也逐漸增加，在滑動窗口大小設為16時達到最佳的性能，然後隨之下降。因此使用整個對話歷史並不能幫助模型提高性能。這些結果也為後續研究提供了基礎。我們針對三個輔導者相關的特徵進行了消融實驗，以驗證其對我們模型的有效性。如表7所示，三個特徵中對模型提升最顯著的是Strategy Feature，性能在Macro F1上提升了約2.5%。雖然Orientation Feature的提升不如Strategy Feature明顯，但在Macro F1上還是帶來了約0.5%的提升。此外Position Feature的加入都進一步提高了模型的性能。而我們的模型在結合了這三個特徵後，在Macro F1上提高了近 3.5%。此實驗也結果說明了這三個輔導者相關的特徵在我們的任務中使用的重要性。

表7、三個輔導者相關的特徵比較

**三、成果自評**

本計畫為三年期計畫之最後一年，在本年度的計畫執行過程中，我們已經順利完成生物醫學專名識別、心理健康狀況檢測、自殺風險評估、情緒輔導對話的策略預測共四項研究。成果包含相關研究論文4篇，其中三篇分別被國際期刊BMC bioinformatics、Journal of Intelligent Information Systems、Multimedia Tools and Applications接受。

三年的時間下來，我們深刻體會到醫療和科技技術的結合並不是一件簡單的事，

當中包括資料的處理和最後結果的解釋，都需要不同背景的專業知識和一定經驗的累積，我們也相信還有許多應用尚未被發展出來，但這三年的時間，我們投入了許多時間和心力在分析醫療資料、社群媒體資料和問答網站資料，並運用深度學習技術在各個應用上，例如，出加護病房預測、情緒原因分析、自殺風險評估等等，當中許多成果也被國際期刊所接受，這已經達到了我們最一開始的預期成果，也代表了我們的成果在這方面做出了一定的貢獻。

**已接受、發表之論文**

1. Zhang, Z., & Chen, A. L. (2022). Biomedical named entity recognition with the combined feature attention and fully-shared multi-task learning. BMC bioinformatics, 23(1), 1-21.

2. Lin, Y. S., Tai, L. K., & Chen, A. L. (2023). The detection of mental health conditions by incorporating external knowledge. Journal of Intelligent Information Systems, 1-22.

3. Tsai, Y. S., & Chen, A. L. (2023). Suicide risk assessment using word-level model with dictionary-based risky posts selection. Multimedia Tools and Applications, 1-20.

**碩士畢業論文**

1. Zhi Yu Zhang, “Biomedical Named Entity Recognition with the Combined Feature Attention and Fully-Shared Multi-Task Learning,” National Tsing Hua University, 2022.

2. Yun Sheng Lin, “The Detection of Mental Health Conditions by Incorporating Knowledge Retrieval,” National Tsing Hua University, 2022.

3. Yun Sheng Tsai, “Suicide Risk Assessment using Word-Level Model with Dictionary-Based Risky Posts Selection,” National Tsing Hua University, 2022.

4. Yi Feng Lin, “Predicting the Following Support Strategy during the Emotional Support Dialogue by Integrating the Dialogue History and Supporter Features,” National Tsing Hua University, 2022.

**參考文獻**

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv*, p. preprint arXiv:1810.04805.

[2] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, (p. 32).

[3] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., & ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv*, p. preprint arXiv:1907.11692.

[4] Lee, J., Yoon, W., Kim, S., Kim, D. K., So, C. H., & Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, pp. 36(4), 1234-1240.

[5] Tian, Y., Shen, W., Song, Y., & Xia, F. (2020). Improving biomedical named entity recognition with syntactic information. *BMC bioinformatics*, pp. 21(1), 1-17.

[6] Miller, A., Fisch, A., Dodge, J., Karimi, A. H., Bordes, A., & Weston, J. (2016). Key-value memory networks for directly reading documents. *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, (pp. 1400-1409).

[7] Tian, Y., Song, Y., Ao, X., Xia, F., Quan, X., Zhang, T., & Wang, Y. (2020). Joint Chinese word segmentation and part-of-speech tagging via two-way attentions of auto-analyzed knowledge. *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, (pp. 8286-8296).

[8] Crichton, G., Pyysalo, S., Chiu, B., & Korhonen, A. (2017). A neural network multi-task learning approach to biomedical named entity recognition. *BMC bioinformatics*, pp. 18(1), 1-14.

[9] Chai, Z., Jin, H., Shi, S., Zhan, S., Zhuo, L., & Yang, Y. (2022). Hierarchical shared transfer learning for biomedical named entity recognition. *BMC bioinformatics*, pp. 23(1), 1-14.

[10] Huang, K., Huang, D., Liu, Z., & Mo, F. (2020, November). A Joint Multiple Criteria Model in Transfer Learning for Cross-domain Chinese Word Segmentation. *In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (pp. 3873-3882).

[11] Arman Cohan, Bart Desmet, Andrew Yates, Luca Soldaini, Sean MacAvaney, and Nazli Goharian. SMHD: a Large-Scale Resource for Exploring Online Language Usage for Multiple Mental Health Conditions. in *Proceedings of the 27th International Conference on Computational Linguistics*. 2018.

[12] Zheng Ping Jiang, Sarah Ita Levitan, Jonathan Zomick, and Julia Hirschberg. Detection of mental health from reddit via deep contextualized representations. in *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis*. 2020.

[13] Ankit Murarka, Balaji Radhakrishnan, and Sushma Ravichandran. Classification of mental illnesses on social media using RoBERTa. in *Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis*. 2021.

[14] Tao Gui, Liang Zhu, Qi Zhang, Minlong Peng, Xu Zhou, Keyu Ding, and Zhigang Chen. Cooperative multimodal approach to depression detection in twitter. in *Proceedings of the AAAI conference on artificial intelligence*. 2019.

[15] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, and Wenwu Zhu. Depression detection via harvesting social media: A multimodal dictionary learning solution. in *IJCAI*. 2017.

[16] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. in *Seventh international AAAI conference on weblogs and social media*. 2013.

[17] Glen Coppersmith, Mark Dredze, and Craig Harman. Quantifying mental health signals in Twitter. in *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality*. 2014.

[18] Andrew G Reece and Christopher M Danforth, Instagram photos reveal predictive markers of depression*.* EPJ Data Science, 2017. **6**(1): p. 15.

[19] Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. A knowledge-grounded neural conversation model. in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018.

[20] Dongfang Li, Baotian Hu, Qingcai Chen, Weihua Peng, and Anqi Wang. Towards medical machine reading comprehension with structural knowledge and plain text. in *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*. 2020.

[21] Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. Dense Passage Retrieval for Open-Domain Question Answering. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.

[22]Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, and Tim Rocktäschel, Retrieval-augmented generation for knowledge-intensive nlp tasks*.* Advances in Neural Information Processing Systems, 2020. **33**: p. 9459-9474.

[23] Shing, H.-C., Nair, S., Zirikly, A., Friedenberg, M., Daumé III, H., & Resnik, P. (2018). Expert, crowdsourced, and machine assessment of suicide risk via online postings. Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic.

[24] Matero, M., Idnani, A., Son, Y., Giorgi, S., Vu, H., Zamani, M., Limbachiya, P., Guntuku, S. C., & Schwartz, H. A. (2019). Suicide risk assessment with multi-level dual-context language and BERT. Proceedings of the sixth workshop on computational linguistics and clinical psychology.

[25] Gaur, M., Alambo, A., Sain, J. P., Kursuncu, U., Thirunarayan, K., Kavuluru, R., Sheth, A., Welton, R., & Pathak, J. (2019). Knowledge-aware assessment of severity of suicide risk for early intervention. The world wide web conference.

[26] Sawhney, R., Joshi, H., Gandhi, S., & Shah, R. R. (2021). Towards ordinal suicide ideation detection on social media. Proceedings of the 14th ACM International Conference on Web Search and Data Mining.

[27] Sawhney, R., Neerkaje, A. T., & Gaur, M. (2022). A Risk-Averse Mechanism for Suicidality Assessment on Social Media. *Association for Computational Linguistics 2022 (ACL 2022)*.

[28] Zhou, H., et al. Emotional chatting machine: Emotional conversation generation with internal and external memory. in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018.

[29] Zhou, X. and W.Y. Wang. MojiTalk: Generating Emotional Responses at Scale. in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2018.

[30] Ghosal, D., et al. DialogueGCN: A Graph Convolutional Neural Network for Emotion Recognition in Conversation. in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 2019.

[31] Sharma, A., et al. A Computational Approach to Understanding Empathy Expressed in Text-Based Mental Health Support. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.

[32] Majumder, N., et al. MIME: MIMicking Emotions for Empathetic Response Generation. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.

[33] O'Leary, K., et al. “Suddenly, we got to become therapists for each other” Designing Peer Support Chats for Mental Health. in *Proceedings of the 2018*

[34] Hill, C.E., K.M. O'Brien, and K. O'Brien, Helping skills: Facilitating exploration, insight, and action. 1999: American Psychological Association Washington, DC.

[35] Sun, H., et al. PsyQA: A Chinese Dataset for Generating Long Counseling Text for Mental Health Support. in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. 2021.

[36] Liu, S., et al. Towards Emotional Support Dialog Systems. in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 2021.

[37] Peng, W., et al., Control Globally, Understand Locally: A Global-to-Local Hierarchical Graph Network for Emotional Support Conversation*.* arXiv preprint arXiv:2204.12749, 2022.

[38] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014, June). The Stanford CoreNLP natural language processing toolkit. *In Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, (pp. 55-60).

[39] World Health Organization, The WHO special initiative for mental health (2019-2023): universal health coverage for mental health. 2019, World Health Organization.

[40] World Health Organization, Comprehensive mental health action plan 2013–2030. 2021, World Health Organization,, .

[41] Z. Steel, C. Marnane, C. Iranpour, T. Chey, J. W. Jackson, V. Patel, and D. Silove, The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013*.* Int J Epidemiol, 2014. **43**(2): p. 476-93.

[42] V. Patel, S. Saxena, C. Lund, G. Thornicroft, F. Baingana, P. Bolton, D. Chisholm, P. Y. Collins, J. L. Cooper, J. Eaton, H. Herrman, M. M. Herzallah, Y. Huang, M. J. D. Jordans, A. Kleinman, M. E. Medina-Mora, E. Morgan, U. Niaz, O. Omigbodun, M. Prince, A. Rahman, B. Saraceno, B. K. Sarkar, M. De Silva, I. Singh, D. J. Stein, C. Sunkel, and J. UnUtzer, The Lancet Commission on global mental health and sustainable development*.* Lancet, 2018. **392**(10157): p. 1553-1598.

[43] V. Bird, P. Premkumar, T. Kendall, C. Whittington, J. Mitchell, and E. Kuipers, Early intervention services, cognitive-behavioural therapy and family intervention in early psychosis: systematic review*.* Br J Psychiatry, 2010. **197**(5): p. 350-6.

[44] J. Treasure and G. Russell, The case for early intervention in anorexia nervosa: theoretical exploration of maintaining factors*.* Br J Psychiatry, 2011. **199**(1): p. 5-7.

[45] Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead. From ADHD to SAD: Analyzing the language of mental health on Twitter through self-reported diagnoses. in *Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality*. 2015.

[46] Sean Macavaney, Anjali Mittu, Glen Coppersmith, Jeff Leintz, and Philip Resnik. Community-level research on suicidality prediction in a secure environment: Overview of the CLPsych 2021 shared task. in *Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology: Improving Access*. 2021.

[47] Adrian Benton, Margaret Mitchell, and Dirk Hovy. Multitask Learning for Mental Health Conditions with Limited Social Media Data. in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. 2017.

[48] James N Butcher, John R Graham, Yossef S Ben-Porath, Auke Tellegen, WG Dahlstrom, and Beverly Kaemmer, MMPI-2: Manual for administration, scoring, and interpretation (Rev. ed.)*.* Minneapolis, MN: University of Minnesota, 2001.

[49] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders. 5th ed. 2013, Washington, DC.

[50] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, Attention Is All You Need*.* Advances in Neural Information Processing Systems 30 (Nips 2017), 2017. **30**.

[51] Ivan Sekulic and Michael Strube, Adapting Deep Learning Methods for Mental Health Prediction on Social Media*.* W-NUT 2019, 2019. **14**(162.2): p. 322.

[52]李明濱. (2020). *109年自殺防治年報*. <https://www.tsos.org.tw/media/4591#doc-tabs-detail>

[53] Cao, L., Zhang, H., Feng, L., Wei, Z., Wang, X., Li, N., & He, X. (2019). Latent suicide risk detection on microblog via suicide-oriented word embeddings and layered attention. *arXiv preprint arXiv:1910.12038*.

[54] Olfson, M., Building the mental health workforce capacity needed to treat adults with serious mental illnesses. Health Affairs, 2016. 35(6): p. 983-990.

## 111年度專題研究計畫成果彙整表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **計畫主持人：**陳良弼 | | | **計畫編號：**109-2221-E-468-014-MY3 | | |
| **計畫名稱：**基於多元資料及深度學習技術之醫療服務應用 | | | | | |
| 成果項目 | | | 量化 | 單位 | 質化  （說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等） |
| 國內 | 學術性論文 | 期刊論文 | 0 | 篇 |  |
| 研討會論文 | 0 |  |
| 專書 | 4 | 本 | 1. Zhi Yu Zhang, “Biomedical Named Entity Recognition with the Combined Feature Attention and Fully-Shared Multi-Task Learning,” National Tsing Hua University, 2022. 2. Yun Sheng Lin, “The Detection of Mental Health Conditions by Incorporating Knowledge Retrieval,” National Tsing Hua University, 2022. 3. Yun Sheng Tsai, “Suicide Risk Assessment using Word-Level Model with Dictionary-Based Risky Posts Selection,” National Tsing Hua University, 2022. 4. Yi Feng Lin, “Predicting the Following Support Strategy during the Emotional Support Dialogue by Integrating the Dialogue History and Supporter Features,” National Tsing Hua University, 2022. |
| 專書論文 | 0 | 章 |  |
| 技術報告 | 0 | 篇 |  |
| 其他 | 0 | 篇 |  |
| 國外 | 學術性論文 | 期刊論文 | 3 | 篇 | 1. Zhang, Z., & Chen, A. L. (2022). Biomedical named entity recognition with the combined feature attention and fully-shared multi-task learning. BMC bioinformatics, 23(1), 1-21.. 2. Lin, Y. S., Tai, L. K., & Chen, A. L. (2023). The detection of mental health conditions by incorporating external knowledge. Journal of Intelligent Information Systems, 1-22. 3. Tsai, Y. S., & Chen, A. L. (2023). Suicide risk assessment using word-level model with dictionary-based risky posts selection. Multimedia Tools and Applications, 1-20. |
| 研討會論文 | 0 |  |
| 專書 | 0 | 本 |  |
| 專書論文 | 0 | 章 |  |
| 技術報告 | 0 | 篇 |  |
| 其他 | 0 | 篇 |  |
| 參與計畫人力 | 本國籍 | 大專生 | 0 | 人次 |  |
| 碩士生 | 5 | 蔡昀陞、林鎰鋒、林昀昇、嚴翎愷、李昀叡 |
| 博士生 | 0 |  |
| 博士級研究人員 | 0 |  |
| 專任人員 | 0 |  |
| 非本國籍 | 大專生 | 0 |  |
| 碩士生 | 0 |  |
|  |  |  |
| 博士生 | 1 | SYAUKI AULIA THAMRIN |
| 博士級研究人員 | 0 |  |
| 專任人員 | 0 |  |
| 其他成果  （無法以量化表達之成果如辦理學術活動  、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。） | | |  | | |