

# 應用自然語言技術於非營利組織捐款留言 之意圖分類

## 摘要

非營利組織在當代社會中扮演著極其重要的角色，彌補了政府與企業無法覆蓋的範疇，組織募集的資金主要用於推動各項計畫、倡議和人道救援，然而，要持續成功地籌措資金，有效的募款策略和與捐款者溝通使其持續捐款，成為其成功與長久的關鍵。本研究目的在於運用先進的自然語言處理技術，對特定非營利組織平台的捐款留言數據進行自動化分類和分析，提高非營利組織對捐款留言的分析效率和精確度，從而更全面地理解捐款者的意圖。

本研究創建一個包含 10 種精細定義類別標籤的分類器，這些標籤不僅涵蓋捐款者的情感表達，也明確描繪出捐款的具體意圖，使非營利組織能深入理解捐款者的動機，從情感支持「關愛與祝福」到具體的財務貢獻「稅務」，每個標籤都精確地對應捐款者的特定需求和期望，如「同理認同」和「環保」標籤反映了捐款者對組織使命的支持或對特定事業的貢獻意願，通過了解捐款者的目的以及他們表達的多樣情感和意圖，組織能更深入地洞察捐款者心中的藍圖，這有助於精緻地溝通並維繫既有捐款者關係，同時對潛在捐款人制定更精準的募款策略，從而提升非營利組織的募款效率。

在研究過程中，使用深度學習模型 BERT，並採用 SVM、XGBoost 機器學習算法與 VotingClassifier 集成學習方法。實驗結果顯示，BERT 模型在處理捐款留言的分類任務上表現最佳，達到 87.92% 的分類準確性，大大提升了處理效率和準確性，減少了人力的投入。本研究已完成專屬於非營利組織領域的訓練資料集，為未來相關領域的研究提供了一個寶貴的資源，彌補了原本相關領域中的資料集缺口。

基於本研究的成果，建議未來研究應擴大數據集範圍，引入更多樣化的 NPO 平台和捐款留言，以進一步提高模型的泛化能力和精確性，研究期望為非營利組織提供實用的工具，以優化募款策略和增強捐款者的互動體驗，促進非營利組織的社會貢獻。

**關鍵字：**非營利組織、捐款留言分析、自然語言處理、BERT。

## 第一章 緒論

### 第一節 研究背景與動機

在現今社會，非營利組織（Nonprofit Organization, NPO）扮演著填補政府與企業無法覆蓋的領域的角色，緩解社會問題、促進社會和諧與發展，NPO 也肩負著對社會更大的使命與責任，其運作資金多數來自社會大眾的捐款，形成了所謂的「公共財」，用來推動各項計畫、倡議與人道救援工作。然而，傳統的募款方式如報章雜誌、廣播、電視媒體等，已無法滿足現今快速變化的社會需求。特別是 2020 年 COVID-19 大流行後，數位行銷和多樣化的募款方式成為趨勢，但這些新興渠道難以精確了解捐款者的具體意圖。

本研究選擇分析 NPO 官方網站上捐款留言，透過自然語言處理技術，進一步了解捐款者的意圖和需求。這樣的分析能夠幫助 NPO 制定更具前瞻性的募款策略，提升募款效率與組織的可持續發展能力，並增進與捐款者之間的互動和溝通。

### 第二節 研究目的

#### 一、開發自動分類捐款留言的工具

在資訊爆炸的今天，手動處理捐款留言已難以應付。因此，本研究目的是運用人工智慧技術，開發一個自動分類捐款留言的工具，精準捕捉捐款者的關注點、需求和意見，提升 NPO 的募款策略和溝通方式。

#### 二、資料前處理與多標籤分類

對捐款留言進行資料前處理，去除雜訊，並進行 10 大類別的多標籤分類，利用 BERT 預訓練模型和 K-fold 交叉驗證方法來評估模型效能。

#### 三、提升 NPO 的募款效能與社會影響力

透過高效處理和分析大量捐款留言，本研究為 NPO 提供關鍵見解，提升組織的募款效能和社會影響力。

#### 四、建立捐款者留言之意圖資料集

建立一個豐富的捐款者留言意圖資料集，為未來模型訓練和優化提供基礎，幫助 NPO 更有效地理解和回應捐款者需求，制定精確的募款策略。

### 第三節 研究流程

本研究旨在探討 NPO 捐款留言的自動分類及其意圖分為五個章節：

緒論：介紹研究背景、動機和目的。文獻探討：回顧 NPO 捐款相關研究、留言分析應用、深度與機器學習在情感分析中的應用。研究方法：詳述研究架構、程序、人工標記、分析步驟和評估指標。實驗設計與結果：展示數據處理結果、模型表現評估與錯誤分析。結果與建議：總結研究結論貢獻、限制和未來研究方向。

## 第二章 文獻探討

### 第一節：非營利組織捐款的相關研究

近兩年，COVID-19 全球大流行對慈善公益領域帶來巨大衝擊，台灣亦未能倖免。後疫情時代，提升捐款持續捐贈頻率對非營利組織（NPO）至關重要。Montgomery (2012) 指出，募款每增加 1%，捐款將減少 0.84%，因此，降低流失率並鼓勵現有捐款人持續捐贈，對組織的長期穩定發展具有重要作用。

捐款行為受多種因素影響，包括年齡、收入、性別、地區和教育程度。研究顯示，個體經濟狀況是一個關鍵因素，捐款額度與捐款者年收入正相關，意味著高收入戶的捐款金額較高（王郁絮，2023）。此外，捐款總金額在性別上存在差異，儘管男性捐款總額較高，但女性捐款意願與頻率較高。陳婉玲（2014）認為，性別差異為募款策略提供了深入研究的空間，未來募款策略可以更有針對性地強化對女性的宣傳，以提高她們的參與度。

捐款的常見動機包括親友介紹、廣告宣傳、節稅、純粹做善事和贈品回饋等。楊雀（2023）指出，同理心觸發利他動機，捐款行為帶來的心理情感豐富了個體的正面情緒。蕭丞勛（2020）亦認為，捐款便利性與捐款意願有顯著正相關，強調多元化捐款方式與管道的重要性。隨著社會進步和科技發展，NPO 需要多元化的捐款方式與管道以適應變化。Kwak（2023）顯示，多元且自由的捐款管道能提高捐款意願與金額，豐富組織財政。後疫情時代，網路行銷興起，各捐款通路的廣告宣傳融合了傳統行銷元素。

台灣社會逐漸邁入老齡化，舊有捐款人的年齡增長，形成捐款主力群體集中在中高齡者，盧憶慧與林清壽（2013）指出，年齡越長，參與捐款比例越高。隨世代改變，捐款方式也逐漸轉移，薛妙賢（2014）指出，利用贈品回饋激勵捐款者，結合企業合作商品進行公益行銷，不僅達到募款效果，也強化企業的社會企業責任（CSR）。

本研究透過對線上捐款者留言的深入分析，設計了 10 種類別分類，包括續捐率、Z 世代捐款禮、稅務和捐款者認同等。NPO 面對後疫情時代的挑戰，需要吸引新捐款者並保持舊有捐款者參與度。傳統與網路捐款方式各有優勢與局限性，但隨著新型態募款手法的出現，NPO 有機會創新吸引捐款者，提升募款效能。

## 第二節 留言分析在社群媒體中的應用

隨著網路技術的迅速發展，社群媒體已成為人們生活中不可或缺的資訊交流平台。Facebook、LINE、Instagram 和 YouTube 等平台如雨後春筍般湧現，成為各類機構與受眾溝通的主要途徑。Facebook 的留言對非營利組織而言，有效運用 Facebook 粉絲專頁能顯著提升與支持者的關係（陳尊鈺，2011），但其效果評估和管理仍需進一步改進（張育嘉，2022）。

LINE 作為廣泛應用的通訊平台，不僅深入到個人生活中，也在政府、企業和非營利組織中發揮著關鍵作用。鄭淑華等（2019）的研究顯示，LINE 的易用性和科技適配性提升了護理人員的溝通效率，但也需要注意隱私保護。衛生福利部的 LINE 官方帳號在疫情期間以「總柴」吉祥物為媒介，有效地傳達疫情資訊和健保政策，展示了 LINE 在時效性和便利性方面的重要性。張蔭淙（2020）的研究進一步強調，LINE 群組在非營利組織中扮演了凝聚成員和快速傳遞訊息的關鍵角色。然而，科技特性帶來的高互動性也可能導致溝通失焦和訊息洗版等問題，因此在 LINE 的運用中必須平衡功能性和隱私保護，以實現最大效能。

Instagram 則作為整合圖片和短影音的平台，逐漸受到年輕族群的青睞。葉慈文（2023）指出，Instagram 不僅讓年輕人更容易接觸數位化的資訊，也促使

推出互動式毒品防制遊戲以提升警覺性。COVID-19 疫情期間，社交媒體成為捐助者的主要依賴平台，非營利組織因此在 Instagram 上成功推行市場營銷策略，以真實的媒體內容增進觀眾連結感，並利用多媒體廣告吸引捐款者 (Cobos, 2022)。

YouTube 作為主要的影音平台，不僅成為現代人日常娛樂和知識吸收的重要途徑，也在疫情期間成為衛生福利部疾病管制署直播聊天室的熱門頻道(孫嘉君, 2023)。隨著疫情教育方式的轉變，YouTube 也成為教學的重要平台之一，蔡瑀 (2023)提出使用 K-means 演算法分類留言主題，協助教育工作者理解觀眾需求，優化教學內容。

社群媒體平台各有其獨特的留言特性和規範，對政府、企業和非營利組織而言，是重要的溝通工具，可協助形象建構和宣傳目標的達成。然而，在使用這些平台時，必須了解它們的特性和限制，以及遵守相應的主題和留言規範。因此，本研究透過官方網站線上捐款的留言分析，提出了 10 種類別，旨在深入探討捐款者的意圖，並為非營利組織的募款策略提供更精準的導向。這一方法希望能夠促進募款活動的順利進行，並有效提升捐款效率，從而更好地服務社會大眾。

### 第三節 深度學習和機器學習在情感分析的應用

近年來，深度學習和機器學習技術的迅速發展對情感分析和文本分類領域帶來了顯著影響，特別是在社群媒體等領域，這些技術不僅能夠深入挖掘大量數據中的情感和意圖，還能夠幫助 NPO 更好地了解社會大眾的需求，改進組織的溝通策略，提升公共服務的效率和質量。

在情感分析方面，BERT (Bidirectional Encoder Representations from Transformers) 由 Google 於 2018 年發表，採用 Transformer 架構，能夠雙向理解單詞的上下文信息 (Devlin et al., 2018)。王柏淞 (2023) 基於 BERT 模型，研究了社群媒體情緒波動對台灣股市的影響，發現情緒波動可以在長期內影響股價的波動。

除了深度學習，機器學習演算法如支持向量機 (SVM)、eXtreme Gradient Boosting (XGBoost) 和 VotingClassifier 也在情感分析和文本分類中顯示出卓越能力。SVM 是一強大的監督式學習模型，在多領域中被廣泛應用，有效處理了分類和回歸分析的問題 (Cortes & Vapnik, 1995)，且林芝儀 (2020) 研究在金融領域中效果顯著。Huang & Wang (2023) 的研究使用 XGBoost 中國新浪微博的貼文數據集進行情感分析，探討 COVID-19 大流行時對中國民眾心理健康的影響和 Singh 等人 (2023) 利用 VotingClassifier 通過在複雜數據集中提升模型性能，展現了在不同應用場景中的重要作用，在肺癌分類高精準確性和效率。

根據文獻探討，深度學習和機器學習在 NLP 領域展現了潛力和效能。這些技術能準確洞察人們的情感和行為，廣泛應用於政經、學術、法律等各領域，為決策提供重要參考。本研究旨在開發捐款留言分類器，將留言分為 10 種類別，以精確分析捐款者意圖，不同於傳統情緒分析的方式。該工具預期能快速準確地分析捐款留言，為組織的募款策略提供支持，精確捕捉社會大眾的情感意圖。

### 第三章 研究方法

#### 第一節 研究架構

本研究核心目標是針對捐款網站上的非結構化捐款留言進行深入分析，包括資料預處理與人工標記，以建立一個適合分析的資料集，接著，利用深度學習技術建立一個模型，專注於對這些留言進行多標籤單類別的二元分類預測，以期精準捕捉捐款者的意圖。

#### 第二節 研究程序

##### 一、資料蒐集

本研究蒐集 NPO 官方網站捐款資料庫中的捐款留言，這些留言反映了捐款者在捐款過程中留下的評論、反饋與需要，通常包含了對組織、活動、服務計畫的支持與感謝，和其他需求與問題。留言的內容與長度各不相同，有中文、英文及表情符號等，有些可能只是簡短的感謝語，有些較為詳細和情感豐富的反饋，本研究涵蓋的捐款留言來自 2020 年 1 月到 2023 年 7 月之間的所有捐款者。

##### 二、資料預處理

預處理資料是確保模型訓練前資料品質良好、特徵擷取一致性的重要步驟。雖然在一般研究中，Emoji 表情符號和特殊符通常帶有情感意圖，但為了保持留言分析的一致性和清晰性，當它們已跟文本留言的情感表達重複時，這些符號在 NPO 留言中的意義通常不超過文本本身的意思。去除這些非文字元素可以幫助我們更精確地評估留言的情感色彩，而不受外部符號的干擾。本研究在收集 NPO 的捐款留言後，首先進行人工處理，檢視每條留言，確保刪除了所有無關字符、重複內容和特殊情況，這些步驟不僅增加了資料一致性，還確保了最終模型在捐款留言分類中的準確性和效能，將進行以下幾項資料雜訊清理工作：

1. Emoji 表情符號與特殊符號：在本資料集類別分類中屬重複或無改變語意。
2. HTML 中的特殊標籤與語法：為避免留言造成留言斷句或語意不同，如<br/>。
3. 標點符號：標點符號的有無不影響捐款留言原意。
4. 重複留言：在模型訓練中重複資料會影響訓練效果，增加模型負擔。

經過上述處理後，形成本研究的資料集，其中留言的字數範圍為 2 到 254 個字數，下面表格為資料集處理前後的範例。

原始捐款留言	處理後捐款留言
祝福每一位山上的孩子/朋友  南無阿彌陀佛💕！	祝福每一位山上的孩子朋友南無阿彌陀佛
愛心的力量無限，願我們一起為公益 出一份微薄之力。(*´ ˘ `*)♡	愛心的力量無限願我們一起為公益出一份微薄之力

#### 第三節 人工標記

本研究建立在文獻探討的基礎上，針對捐款留言進行分析，設計了一個捐款意圖分類器，包含 10 個類別標籤：「再次支持」、「環保」、「稅務」、「贈品」、「個資」、「同理認同」、「希望」、「感謝」、「加油與期待」及「關愛與祝福」，以反映捐款者的具體意圖、情感表達和問題。這些類別用來識別捐款者對持續支持、環

保、稅務與捐款贈品的具體需求，並揭示捐款者的內在動機和對組織的認同感。

透過這一分類器，非營利組織能更全面理解捐款者的需求，制定更有效的溝通策略和募款活動，提高捐款者參與度和捐款效率，增強與捐款者的連接，為組織帶來更穩定的支持基礎。而為提升人工標記的效率和準確性，需要細化每個類別的定義，提供具體的標記指南和範例。這些指南和範例明確界定每個類別的範疇和情境，有助於標記者快速準確地識別和分類捐款留言，提升資料品質和後續分析的有效性。以下是類別的標記準則說明與範例。

一、「再次支持」類別，捐款人表示他們願意再次捐款給慈善機構，這可能是因為他們對機構的工作印象深刻，與更改信用卡、更新卡號重新授權等授權相關。範例：「我以前捐過款，並且願意繼續支持。」、「我相信這個機構，我會繼續捐款，以幫助更多的孩子。」或「捐款編號、直接打出數字編號、老朋友」。

二、「環保」類別，捐款人表示對於環保的實際行動和期望，通過數位化的方式表達對環境保護的支持，希望將收據、報告、感謝卡等轉為電子化，減少紙張的浪費和對環境的負擔。

範例：「不再需要紙本報告」、「收據電子化」或「減少砍樹、電子通知」。

三、「個資」類別，捐款人關注的隱私保護，特別是個人資料和金融信息安全，以防止資料外洩或遭到不當使用，及捐款資訊是否會被公布。

範例：「我的個人資訊會被妥善保護嗎？」或「請確保我的個人資訊不會外洩」

四、「稅務」類別，捐款人通過合法途徑支持慈善活動，並在稅務上獲得相應的減免，提供公司名、公司統編、個人身分證字號等獲得相關的稅務收據。

範例：「請幫忙上傳國稅局，不需收據」、「提供公司統編與個人身分證字號」。

五、「贈品」類別，機構會提供捐款滿額禮贈品作為回饋，以表達對民眾慷慨捐助的感謝之意，會收到紀念品、書籍或其他小禮品。

範例：「想索取感謝禮」或「不需要任何贈品」。

六、「感謝」類別，捐款者表達肯定機構的努力與慈善活動成效的認可，與工作的贊賞和對偏鄉和弱勢族群的付出表示讚揚，並分享他們對機構的深刻印象。

例如：「我由衷感謝你們的工作，你們做得太棒了！」或「謝謝你們為這些偏鄉孩子所做的一切」。

七、「加油與期待」類別，捐款人表達對孩童與慈善機構的支持和期待，期待捐款能有效幫助機構推動特定計畫與對特定族群的鼓勵。

範例：「我期待著看到這個計畫的持續發展。」或「您們辛苦了！加油！」。

八、「關愛與祝福」類別，捐款人對弱勢群體的深切關懷與良好祝願，能夠讓弱勢族群能夠過上幸福生活的深切愛意、關愛之情。

範例：「我由衷地愛護這些孩子讓他們感受到關愛和關懷，讓愛心溫暖他們的心」。

九、「希望」類別，捐款人對弱勢群體的支持與鼓勵，希望能夠給予他們力量和動力，追求自己的夢想，取得成功，並祝願他們能夠迎來更美好的生活。

範例：「我希望這些孩子能夠實現他們的夢想，走向成功的道路。」

十、「同理認同」類別，捐款人對處於不利境地的弱勢群體的深刻同理心與認同

感，他們分享自己在生活中遇到困難時的經歷並表示他們不孤單。此外，他們也會表達對慈善機構發展計畫的認同、贊同和支持。

例如：「我也曾經面臨過困難，我明白他們的處境」、「響應某計畫」。

在人工標記過程中，標記人員需充分理解類別定義，以確保準確性和高效率。鑒於一則捐款留言可能包含多個類別，本研究採用多標籤標記方式，使用預先定義的十種類別進行標記。對於單一類別的留言，例如「謝謝你們的每一份真心」，標記為感謝類別；對於多類別的留言，例如「收據開叻應實業股份有限公司 28586361 金額和叻應的捐款合併就好」，標記為個資、稅務類別；對於三種類別的留言，例如「我的編號是 5245 收劇請開公司名稱群之囑傳播股份有限公司」，標記為再次支持、稅務和個資類別，如下表說明。

留言	標記類別
謝謝你們的每一份真心	感謝
收據開叻應實業股份有限公司 28586361 金額和叻應的捐款合併就好	個資、稅務
我的編號是 5245 收劇請開公司名稱群之囑傳播股份有限公司	再次支持、稅務、個資

#### 第四節 分析步驟

近年來，隨著深度學習技術的快速發展，基於神經網絡的 NLP 模型，尤其是 BERT 模型，已在眾多 NLP 任務中取得了顯著的成功。在本研究中，利用預訓練的 BERT 模型和監督式學習方法對捐款留言數據進行分類，建立一個捐款意圖的分類器，對捐款留言進行分析，使用 BERT 的中文預訓練模型「bert-base-chinese」，經由不同的 LEARNING RATE 和 EPOCHS 的組合，其餘參數皆保持一致，最後在對類別評估結果進行分析，其模型規畫說明如下。

##### 一、數據處理

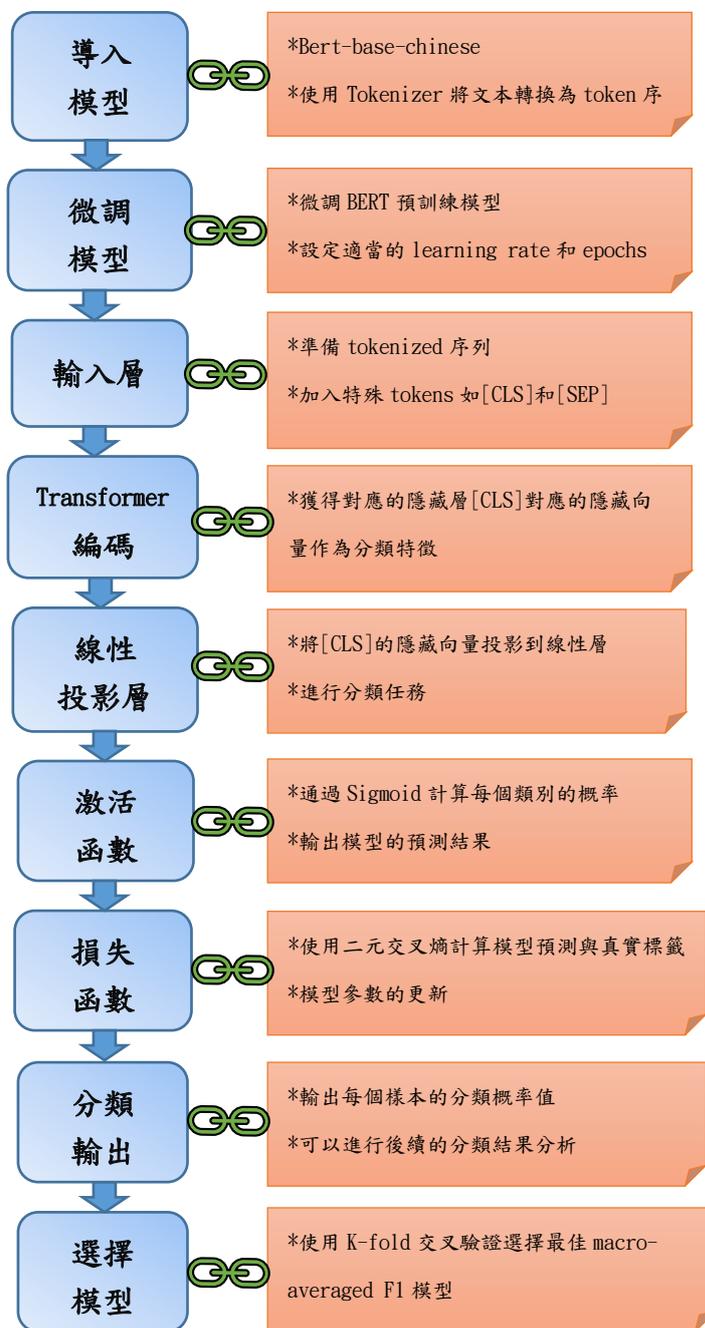
1. 將已經標籤好的資料集作為模型訓練的輸入值，確保數據質量與一致性
2. 對每一個類別訓練獨立的二元分類模型

##### 二、模型訓練

1. 導入基於 BERT 架構的預訓練序列 BertForSequenceClassification 二分類模型和相對應的 Tokenizer，以進行序列的預處理。
2. 微調 (fine-tuning) BERT 預訓練模型，設定適當的 LEARNING RATE 與 EPOCHS。
3. 輸入層 (Input Layer) 將留言序列 token 化表示，每則留言都是一個輸入樣本。
4. 輸入序列通過 Transformer Encoder 編碼，獲得對應的隱藏層 [CLS] 對應的隱藏向量作為分類特徵。
5. 線性投射層-取 [CLS] 輸入到線性層，產出 logits 向量。
6. 激活層-logits 向量經由 Sigmoid 激活函數，獲得的概率值作為最終輸出。
7. 損失函數-使用了二元交叉熵損失函數，計算模型輸出與真實標籤之間的誤差，以指導模型學習。
8. 分類結果 (Classification Output)-輸出概率值。

9. 模型選擇-使用 K=5 的 K-fold 交叉驗證，選擇驗證集的 macro-averaged F1 分數最好的一個 fold 模型作為最優模型。

10. 下圖為模型訓練流程圖。



### 三、模型評估

1. 針對於每個類別，利用訓練的模型對測試集數據進行預測。
2. 最終以 Macro-average F1-score 為評估指標。

### 四、實驗結果儲存

1. 將實驗結果保存到 Excel 文件中，以利後續分析。
2. 將訓練好的模型儲存，以便未來使用。

## 第五節 評估指標

本研究採用將多標籤分類問題轉化為多個獨立的二元分類任務，對每一個分類類別(如再次支持、環保等)，建立一個獨立的二元分類模型，這樣對於 10 個分類類別，需要分別訓練 10 個獨立的二元 BERT 分類器模型，每個類別獨立評估。將每個類別獨立訓練預估，本研究採用 Macro-average F1-score 衡量指標，來評估模型的效能，混淆矩陣說明如下面表格，衡量指標的說明如下。

類別		實際	
		有標記	沒標記
預測	有標記	TP (True Positive) 預測正確	FP (False Positive) 預測錯誤
	沒標記	FN (False Negative) 預測錯誤	TN (True Negative) 預測正確

捐款目的與意圖分類器，將 10 個類別分為有標記與沒標記，例如，TP: 正確預測到「每一個類別」有標記的，TN: 正確預測「每一個類別」沒有標記的，FP: 預測錯誤將「每一個類別」沒標記的預測成有標記，FN: 預測錯誤將「每一個類別」有標記的預測為沒標記。

Accuracy 指模型預測「每一個類別」正確預測的比例，Accuracy (i) 是第 i 個類別的 Accuracy，公式如下：

$$Accuracy_i = \frac{(TP + TN)}{(TP + FP + FN + TN)}$$

Precision 指模型正確預測「每一個類別」的精準度，預設有標記實際也是有標記的比例，Precision (i) 是第 i 個類別的 Precision，公式如下：

$$Precision_i = \frac{TP}{(TP + FP)}$$

Recall 是指預測「每一個類別」的召回率，實際有標記中被模型預測到有標記的比例，Recall (i) 是第 i 個類別的 Recall，公式如下：

$$Recall_i = \frac{TP}{(TP + FN)}$$

F1 Score 是指「每一個類別」精確率和召回率的調和平均數，用於綜合二分類模型的性能，F1(i) 是第 i 個類別的 F1 分數，公式如下：

$$F1\ Score_i = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

然而，本研究為多類別分類，我們關注每個類別的性能，以確定模型對於各個類別的預測表現，所以採取 Macro-average F1-score 為衡量指，因為它是計算每個類別的 F1-score，然後將這些分數的平均值作為模型的性能指標，而不考慮類別的樣本數量，讓捐款目的與意圖 10 個類別都被視為同等重要，不考慮

類別之間資料的不平衡問題。

Macro-average F1-score 計算每個類別的 F1 分數，並對所有 F1 取平均值，N 是總類別數，F1(i) 是第 i 個類別的 F1 分數，為精確率和召回率的平均，公式如下：

$$\text{Macro - average F1 - score} = \frac{1}{N} \sum_{i=1}^N \text{F1 - score}_i$$

在實驗設計中，將與其他模型進行比較，並將 Macro-average F1-score 統一作為的最終評估指標。

## 第四章 實驗設計與結果

### 第一節 實驗資料

本研究的資料來自非營利組織捐款網站，涵蓋 2020 年至 2023 年 7 月，共包含 22994 筆留言。經資料前處理後，篩選出 4717 筆留言作為研究資料集。使用標記準則進行人工標記，統計數據顯示：再次支持 942 次、環保 153 次、個資 316 次、稅務 228 次、贈品 269 次、感謝 701 次、加油與期待 720 次、關愛與祝福 611 次、希望 694 次、同理認同 608 次，共標記 5242 次。這反映了捐款者對各類別的關注度，分佈佔比如表 16 所示。由於採用多標籤標記方式，有 502 則留言包含 2 種類別，23 則留言包含 3 種類別，表明捐款留言具有多樣性，涉及多種情感意圖，分類標記佔比與多類別留言的佔比如下面兩個表所示。

分類標籤	標記數量	佔比
再次支持	942	18.0%
環保	153	2.9%
個資	316	6.0%
稅務	228	4.3%
贈品	269	5.1%
感謝	701	13.4%
加油與期待	720	13.7%
關愛與祝福	611	11.7%
希望	694	13.2%
同理認同	608	11.6%
10 類總和	<b>5242</b>	100%

分類標籤	標記留言數量	佔比
含一種類別	4192	88.87%
含二種類別	502	10.64%
含三種類別	23	0.49%
總和	<b>4717</b>	100%

根據前述分類好的類別，透過對每個類別內捐款留言的分析，計算出每則留言的最短字數、最長字數以及字數範圍的留言字數，如在「環保」類別中，留言的最短為 6 字，最長為 96 字，字數範圍為 90 字。在「稅務」類別中，留言的最短為 2 字，最長為 254 字，字數範圍為 252 字。深入分析不同類別的字數分佈發現，「稅務」類別的留言範圍最廣，從 2 到 254 字，顯示捐款者在留言涉及稅務問題時可能會從簡單的詢問到提供詳盡的資料。這種廣泛的字數範圍可能與捐款者需要提供詳細的個人和財務資訊以確保稅務抵扣資格有關。相較之下，其他類別如「環保」展現出不同程度的字數分佈，指示出捐款者可能只需簡短表達支持

或提出具體要求。這些分析結果有助於 NPO 深入理解捐款者的溝通風格和需求，從而調整其溝通策略和捐款表單，以更好地滿足捐款者的期望並提高募款效率。

## 第二節 實驗設計

在本研究中，為了訓練模型的有效性和泛化能力，對標記完畢的資料集採用了 StratifiedKFold 分層交叉驗證方法，將資料集分為五個 Folder，確保每個折疊中的類別分佈與整體資料集保持一致，同時透過設置 random\_state 參數為 42，固定了訓練過程中的隨機性。因此，對於資料集中的每一個類別，均分配了五個獨立的 Folder，每一 Folder 中均包含了訓練與測試用的數據。

為了訓練模型，本研究採用深度學習和機器學習兩種主要方法，下列表格為演算法參數設定對照表。

演算法	參數設定
BERT-(bert-base-chinese)	learning rate 為 [1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6, 1e-7, 1e-8, 1e-9, 1e-10] num_epochs 為 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]
SVM	C 為 [0.1, 1, 10, 100, 1000] kernel 為 [linear, poly, rbf, sigmoid]
XGBoost	learning rate 為 [1e-1, 1e-2, 5e-2] max_depth 為 [3, 5, 7, 9]
VotingClassifier(SVM、XGBoost)	C 為 10、Kernel 為 linear(SVM) learning rate 為 1e-1(XGBoost) max_depth 為 5(XGBoost) n_estimators 為 300(XGBoost) Voting 為 Soft

在深度學習方法中，選擇了 BERT 的「bert-base-chinese」預訓練模型，透過調整 learning rate 與 epoch 參數來構建二元分類模型。另一方面，機器學習方法則利用 Jieba 進行斷詞處理，通過 TF-IDF 技術將文字轉換成特徵向量，再配合 SVM 和 XGBoost 演算法，調整 C、kernel 與 learning rate、max\_depth、n\_estimators 等參數來建立二元分類模型。兩種訓練方法都利用了 K 為 5 分層交叉驗證來進行訓練集的切割，最終，最終採用不考慮類別之間資料的不平衡問題的 Macro-average F1-score 來評估與分析模型效果，來全面分析模型的效能。

在機器學習方法基礎上，為了進一步提升模型的穩定性和準確性，本研究構建了一個集成模型 (VotingClassifier)，此集成模型結合了 SVM 和 XGBoost 兩種演算法，使用 Soft Voting 方式，軟投票需要每個分類器對每個樣本進行預

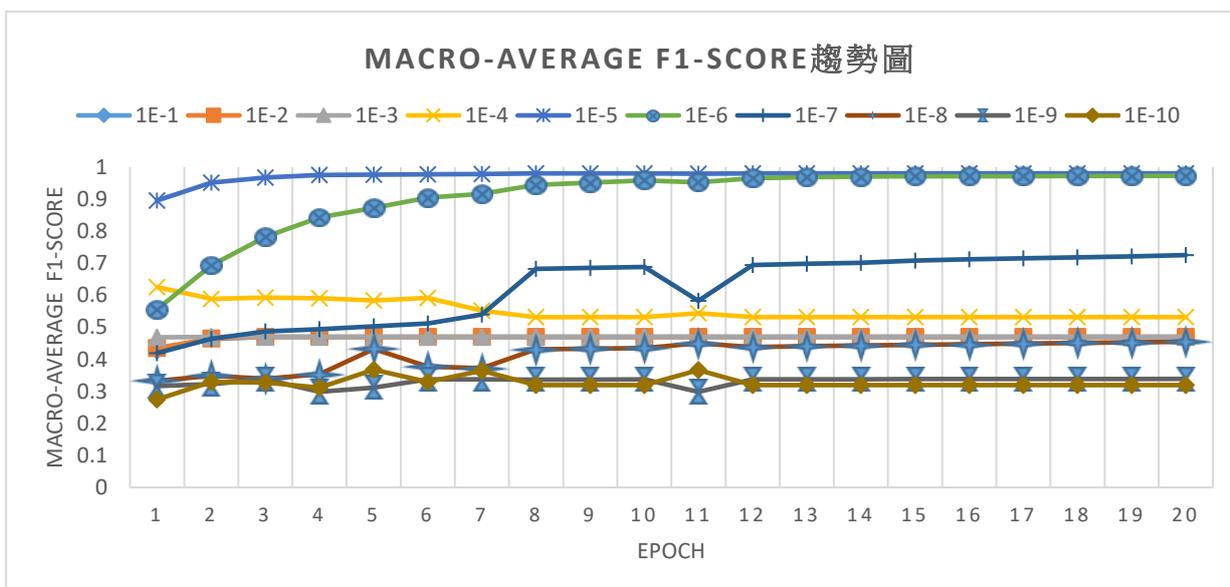
測，並給出該樣本屬於各個類別的概率，然後將這些概率進行平均，以決定最終的預測結果。因為，定義 SVM 時設置了「probability=True」，使其能夠在預測時返回概率分佈，而 XGBoost 則默認支持概率預測，無需額外設置。對於每個模型，會將所有分類器的預測概率進行平均，然後選擇平均後概率最高的類別作為最終預測結果。

### 第三節 實驗結果及分析

本研究旨在透過深度學習與機器學習模型的精細調整，找到效能與效率的最佳平衡。在實驗設計的基礎上，研究顯著提升了模型的預測精確度，增強了結果的穩健性和可信度。接下來將逐步展示各模型在達到最佳效能時的關鍵參數設定，並詳細說明參數優化的過程。

#### 一、BERT 模型結果分析

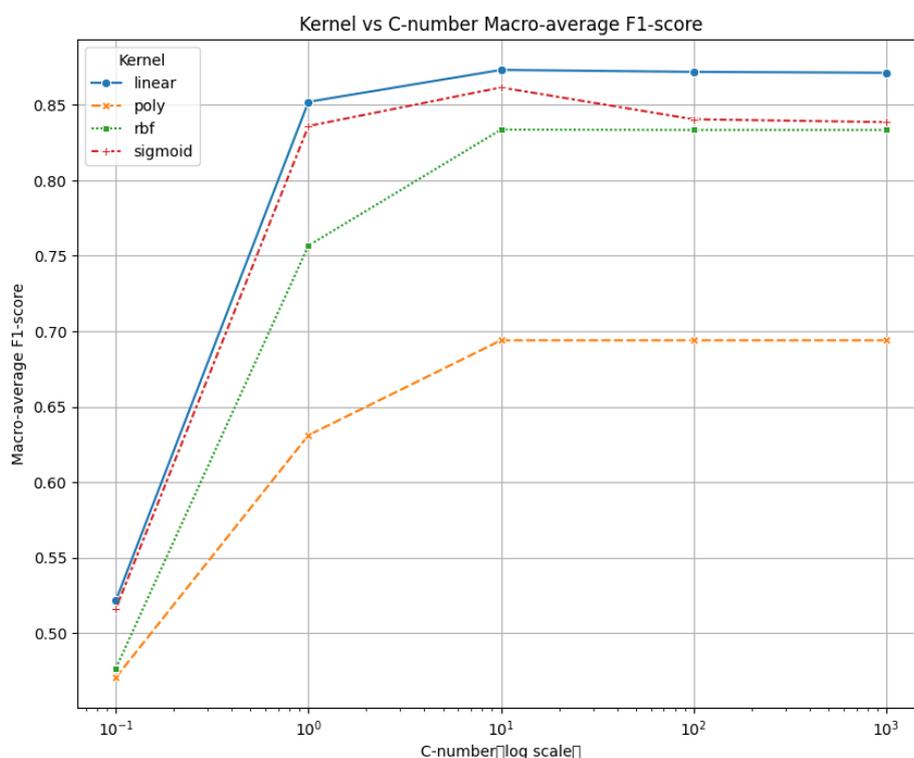
在 BERT 模型的訓練過程中，將 learning rate 設定為  $1e-5$ 、epoch 設定為 13，於驗證集上獲得了最佳的 Macro-average F1-score: 0.98067，參數趨勢圖如下圖所示。由圖中可以發現到，learning rate 在  $1e-5$  的模型表現最好，一開始就領先於其他的模型，隨著 epoch 增加，F1-score 逐漸提高，且在第 13 個 epoch 達到最高，之後趨於穩定並維持在高分數水平。另外 learning rate 在  $1e-6$  的模型起步比較慢，在初始階段的不如  $1e-5$  的模型，但是隨著訓練的進行，其性能呈現出拋物線形狀的上升趨勢，並在第 14 個 epoch 之後趨近  $1e-5$  的 F1-score。然而過高的 learning rate 如  $1e-1$  和  $1e-2$  從一開始就表現不佳，過低的 learning rate 如  $1e-9$  和  $1e-10$  在整個訓練過程中也未能得到較高的 F1-score。在這樣觀察下， $1e-5$  的模型能得到最好效果，因其是一個既不過大也不過小的 learning rate，允許模型快速學習但又不至於因步伐過大而跳過最佳解。



#### 二、SVM 模型結果分析

在 SVM 模型訓練過程中，將 C 設定為 10、Kernel 設定為 linear，於驗證集上獲得了最佳的 Macro-average F1-score: 0.87308，參數趨勢圖如下圖所

示。由圖中可以發現到，在各種 kernel 中，大家起始點分數都相對低，但隨著 C 值增加而改善，而 linear 表現相對的穩定，F1-score 在提升 C 值後很快達到峰值並穩定高分。poly 在低 C 值時的表現較差，但隨著 C 值增加，效能顯著提升，之後在 C=10 趨於穩定。顯示在較高的 C 值下，模型更精確地擬合數據，而在較低的 C 值下，未能捕捉到數據的複雜性。rbf 如同 linear 在 C=1 迅速提高效能，然後達到平穩，但分數未超越 0.85。sigmoid 的表現隨著 C 值增加變化卻是跳躍式改變，呈現 C 值改變不是穩定成長的 F1-score。顯示對於目前的數據集來說，sigmoid 無法經由 C 值提升得到穩定的成效。因此，在這個數據集上 linear 為相對最佳選擇，一開始就取得較高的分數，隨著 C 值增加保持穩定，顯示數據集上，特徵間的關係可能較為線性，且線性決策邊界足以有效分類不同的類別。而其他模型，特別是 poly 和 rbf，對 C 值的改變有相對應的成長，表示找到合適的 C 值對模型效果有顯著的影響。此外，當 C 值足夠高時，所有模型以達到一定的平穩效果，表示模型已經從數據中學習到了足夠多的資訊，進一步增加 C 值也無法對模型增加過多的效能。

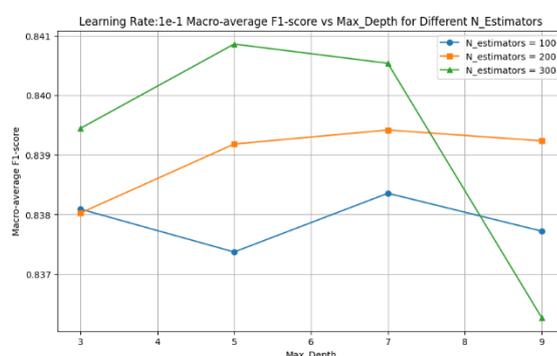
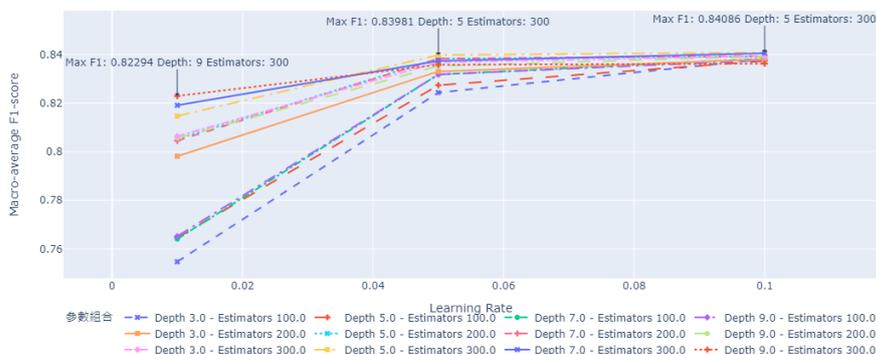


### 三、XGBoost 模型結果分析

在 XGBoost 模型的訓練過程中，將 learning rate 設定為 1e-1、max\_depth 設定為 5、n\_estimators 設定為 300，於驗證集上獲得了最佳的 Macro-average F1-score: 0.84086，參數趨勢圖如下二個圖所示，由圖中可以發現到，隨著 learning rate 的提高，模型性能整體呈現上升趨勢，但在達到 learning rate 為 0.05 的峰值後趨於平穩甚至有輕微的下降，因為過低的學習率使得模型更新緩慢，而過高則可能讓模型更新過快，未能充分學習到數據中的特徵。另一方面，樹的最大深度對 learning rate 的反應有差異顯著。例如，在 learning rate 為

0.1 且樹深度為 5 時，模型達到最高的 F1-score=0.8408，而樹深度為 9 時，較低的學習率 0.01 卻有較低的 F1-score=0.8229。隨著決策樹的數量增加，模型的性能表現出更大的穩定性，尤其是當決策樹的數量為 300 且深度為 5 時，在 learning rate 為 0.1 表現最佳。然而，當模型的深度過高時，即使樹的數量增加，也會導致性能下降，反映出過度擬合的風險。如同，當樹的數量設為 300 時，隨著樹深的增加，F1-score 跟著上升，在樹深度為 5 時達到最大值，但如果深度繼續增加，分數則開始下降，當樹的數量設為 200 也亦同，而樹的數量設為 100 時，隨著樹深的增加，F1-score 隨之上下跳動，這顯示這模型無法捕捉這數據集的特徵，所以在調整 XGBoost 模型時，除了 learning rate 的設定外，必須謹慎平衡樹的深度和數量，以避免不必要的計算負擔和過度擬合，同時保證模型能夠從數據中學習到足夠的信息。

Learning Rate與Macro-average F1-score的關係



#### 四、VotingClassifier 模型結果分析

本研究另設計的 VotingClassifier 採用了 SVM 與 XGBoost 模型的最佳參數配置，期望達到更佳的分類效果。參數將 SVM 的 C 值設定為 10 與 Kernel 設定為 linear; XGBoost 則設置了 learning rate 為 1e-1、max\_depth 為 5、n\_estimators 為 300，為了在複雜的數據集上最大化其性能。實際表現顯示，在測試集上的 Macro-average F1-score 為 0.7945，超越了單獨使用 XGBoost 的表現，其在驗證集上的 Macro-average F1-score 為 0.7965 表現卻未能超越使用 XGBoost 和 SVM 的單一模型。顯示雖然 VotingClassifier 整合了兩種演算法的優勢，但在特定應用場景中，其總體表現仍可能受限於較弱的模型。

通過上述分析，我們可以清晰地看到，在本研究中每個模型的最佳參數組合，因此在 BERT 模型中得到，最好的驗證 Macro-average F1-score 為 0.9806，進行分類預測結果得到 Macro-average F1-score 為 0.8792，是在所有測試的模型中表現最佳，其次是 SVM 模型的驗證 Macro-average F1-score: 0.8731，進行分類預測結果得到 Macro-average F1-score 為 0.8095，與 XGBoost 模型的驗證 Macro-average F1-score 為 0.8408，進行分類預測結果得到 Macro-average F1-score 為 0.76537 與 VotingClassifier 的驗證 Macro-average F1-score 為

0.7965，進行分類預測結果得到 Macro-average F1-score 為 0.7945，如下表所示。

演算法	驗證集	測試集
	Macro-average F1-score	Macro-average F1-score
BERT	0.9806	0.8792
SVM	0.8731	0.8095
XGBoost	0.8408	0.7653
VotingClassifier	0.7965	0.7945

基於實驗結果顯示，BERT 模型展現出卓越的分類效能，通過設定適當的參數 (learning rate:  $1e-5$ 、epoch: 13)，在各類別捐款留言資料中實現了高 Macro-average F1-score。具體成效數據如下表所示，來深入探討其對不同類別的分類效能。在「同理認同」、「加油與期待」等類別，在驗證集上的表現普遍優於測試集，這可能是因為模型在見過的數據上有更好的表現，而在未見過的測試數據上，表現略有下降。特別是在「再次支持」和「贈品」類別中，模型達到了接近完美的分類效果，這表明模型能夠有效捕捉到這些類別特有的語言模式和關鍵信息。對於「環保」和「稅務」等涉及專業知識或具體操作的類別，模型同樣顯示出強大的分類能力，這可能得益於 BERT 的深層語義理解能力。此外，對於如「希望」和「關愛與祝福」這樣的情感導向類別，雖然 F1-score 略低於其他類別，但仍表現出良好的分類效果，顯示了 BERT 在捕捉情感和意圖方面的潛力。總體而言，實驗結果不僅證明了 BERT 在多類別文本分類任務上的有效性，也為未來在類似應用中使用深度學習技術提供了重要參考。這些洞察有助於 NPO 更好地理解 and 運用 AI 技術，以提高數據分析和決策制定的精確性。

類別	驗證集	驗證集	測試集	測試集
	Macro-average F1-score	Macro-average F1-score_std	Macro-average F1-score	Macro-average F1-score_std
個資	0.9816	0.0396	0.8549	0.0537
再次支持	0.9950	0.0126	0.9771	0.0107
加油與期待	0.9606	0.0693	0.7873	0.0580
同理認同	0.9608	0.0760	0.7882	0.0626
希望	0.9796	0.0382	0.8613	0.0543
感謝	0.9896	0.0168	0.9394	0.0193
環保	0.9931	0.0194	0.8939	0.0864
稅務	0.9854	0.0329	0.9256	0.0404
贈品	0.9890	0.0222	0.9400	0.0547
關愛與祝福	0.9714	0.0598	0.8239	0.0506
<b>類別總和</b>	<b>0.9806</b>	<b>0.0387</b>	<b>0.8792</b>	<b>0.0491</b>

## 五、錯誤分析

在前文中，已經詳細說明了 BERT 模型在每個類別上的表現，確定其在本研究中得到最佳的分類效能。然而，本研究並不滿足於僅僅探討它的優勢，而是希望進一步了解其他模型在每個類別上的錯誤預測，通過對這些類別進行更詳細的分析和評估，希望能夠確定導致低效能的具體因素，從而開發更加精確的策略來提高這些類別的預測能力，不僅將提高這些類別的預測精度，也將對整體模型的性能產生正面影響，並為未來的研究提供有價值的參考。

先對每個模型預測錯誤的情況進行了分析，再將每個類別中每個模型都預測錯誤的捐款留言筆數，與有標記沒被預測到 (False Negative, FN) 和沒標記卻預測到 (False Positive, FP) 的筆數整合成一個表格，以進一步了解模型在不同類別上的表現。如表所示，可以清楚地看到哪些類別受到了模型預測錯誤的影響最為顯著，發現「加油與期待」和「同理認同」這兩個類別，在錯誤預測方面表現明顯較差。

類別	預測錯誤	有標記沒被預測(FN)	沒標記卻預測到(FP)
加油與期待	55	50	5
再次支持	11	10	1
同理認同	56	56	0
希望	15	14	1
個資	29	29	0
稅務	9	9	0
感謝	9	6	3
環保	5	5	0
贈品	12	11	1
關愛與祝福	43	42	1
總計	244	232	12

在預測表現較差的「加油與期待」類別中，錯誤預測的留言共同特點是表達支持、鼓勵和未來期望，容易與「感謝」和「希望」等類別混淆。例如，「謝謝您們的用心雖然我能給的不多但是感謝有您們把我們的愛心送給需要的孩子期待他她們未來能成為他人的祝福」和「一起加油疫情期間一定辛苦了」都表達了對未來的期望和支持，這在主觀判斷時容易產生誤差。

「同理認同」類別的錯誤預測留言主要表達對弱勢群體的關懷和志願者的敬意，但常含有抽象或主觀性較強的表達方式，導致模型難以準確理解。例如，「雙薪家庭的緣故很常常都是一個人走路上下課」這樣的描述可能不夠清晰，容易被模型誤解為其他類別的內容。同樣地，一些留言可能包含特定地區或族群的資訊，如「台東阿美族原住民」，這些地域或文化性資訊可能會使模型在分類時產生混淆。

總體而言，模型在處理複雜情感和抽象概念的留言時，因學習機制的局限性和語意理解的主觀差異，容易出現錯誤分類。這反映出當前模型在捕捉和解釋人

類情感豐富性和語言多樣性方面的挑戰，需要進一步優化和確認人工標記的主觀性，以增強其識別細微語意差異的能力。

## 第五章 結果與建議

### 第一節 研究結論與貢獻

本研究透過結合深度學習與機器學習兩種方法來深入分析捐款留言，目的是為了準確掌握捐款者意圖。在深度學習採用 BERT 的「bert-base-chinese」預訓練模型；而在機器學習方面，通過將捐款留言透過 TF-IDF 技術轉化為特徵向量，並結合 SVM、XGBoost 與 VotingClassifier 算法，通過不同參數的設置進行實驗。研究結果顯示，BERT 的「bert-base-chinese」預訓練模型在細微語義捕捉方面表現最佳，為 NPO 在理解捐款者需求與期望方面供有效的分類指標。

由於一些類別含有較多主觀因素、留言涉及多個類別以及類別標記數量不均等問題，導致部分類別的分類效果不佳，本研究雖未提出全新模型或方法，但透過多種演算法與參數的調整加上交叉驗證的方法，仍然使 BERT 模型在平均類別得分上達到了 87.92% 的成效。對於 NPO 來說，希望透過大數據與人工智慧技術分析捐款者留言的類別，以減少人工審閱衍生的錯誤率和分類的時間，提升募款效率和增加募款金額。研究同時完成了一份專屬於 NPO 領域的訓練資料集，為未來相關領域的研究提供了一個寶貴的資源。目前，針對捐款留言分類的研究相對較少，希望本研究的成果和所使用的捐款留言資料集以及分類方法能為未來相關領域的研究提供參考。

### 第二節 研究限制

回顧本研究的實驗過程，我們遇到了幾個關鍵的問題和限制，這些可能影響了研究結果的廣泛使用度和準確度。

一、捐款留言的數量，本研究中使用的捐款留言數量相對較少，對於訓練深度學習模型而言，將會限制捕捉捐款留言多樣性和複雜性的能力，資料量的不足可能削弱模型的泛化能力，進而影響在處理未知數據時的性能表現。

二、人工標記的問題，雖然我們努力制定了一套標準化的分類準則，但由於研究時間限制下僅由單一研究者進行標記，不可避免地帶入了一定的主觀性，可能對資料集的可靠性造成影響。

三、數據不平衡，在本研究蒐集的數據集中，有些類別的樣本較少，導致數據不平衡的問題，這對於訓練機器學習模型而言是一大挑戰，可能讓模型忽略了數量較少的類別。

四、數據來源的局限，本研究主要從特定 NPO 平台蒐集捐款留言，這種選擇可能使得研究結果的適用範圍受到限制，因此在其他平台或組織上所蒐集到的捐款留言可能會有不同的特徵和分類挑戰。

### 第三節 未來研究方向

本研究在處理二元分類問題上展現了潛力，但仍有進一步提升的空間，面對 NPO 實際應用的需求及未來研究的發展，有下列幾個方向可進行擴展和深化。

一、擴增資料集，本研究使用的捐款留言數量有限，限制了模型學習的全面性和

預測的準確性，未來研究應加強從廣泛的 NPO 平台蒐集更多元化的捐款留言，使模型能學習到更豐富的情感表達與層次。

二、分析類別之間的相關性，目前研究著重於獨立類別的分類問題，未來可改用多元標籤分類的評估模型，分析類別間的相互關係，以期通過這些相關性提高分類的精確度和理解的深度。

三、改善人工標記過程，為減少人工標記的主觀性差異，未來的研究需要進一步規範標記流程，引入多位標記者和設計有效的標記指南，增強數據的一致性與可靠性。

四、解決數據不平衡問題，針對少數類別樣本量少的問題，未來研究可使用如生成式對抗網絡 (Generative Adversarial Networks, GANs) 的方法，生成樣本數量較少的類別的新樣本，透過人工增加這些類別的樣本數量，從而緩解數據不平衡的問題以改善模型對少見類別的識別能力。

透過這些方向的改進，未來的研究不僅能突破現有限制，更提升模型的泛化能力和準確性，為 NPO 的募款策略提供更加精準和有效的數據方針，進一步優化與捐款者的溝通和募款成效。

## 參考文獻

- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273-297.
- Cobos, A. (2022). Investigating Instagram Marketing Strategies for Nonprofit Donor Engagement During the Covid-19 Pandemic.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Huang, R., & Wang, X. (2023). Impact of COVID-19 on mental health in China: analysis based on sentiment knowledge enhanced pre-training and XGBoost algorithm. *Frontiers in Public Health*, 11.
- Kwak, J. (2023). The Impact of Donation Method on Donation Decision-Making: Focusing on Implicit Self-theory and priming Effect. Korea International Trade Research Institute.
- Montgomery, A.U. (2012). AN IMPROVED MODEL OF DONATIONS TO NONPROFIT ORGANIZATIONS.
- Singh, O., Singh, K.K., Das, S., Akbari, A.S., & Manap, N.A. (2023). Classification of lung cancer from histopathology Images using a Deep Ensemble Classifier. 2023 IEEE International Conference on Imaging Systems and Techniques (IST), 1-6.
- 王柏淞(2023)。基於 BERT 模型的社群媒體情緒分析對台灣股市長期趨勢的影響探討 [ 碩士論文，國立臺北商業大學資訊與決策科學研究所 ]。  
<https://hdl.handle.net/11296/n89fky>
- 王郁絜(2023)。疫情前後國人捐款支出之分析 [ 碩士論文，國立中央大學產業經濟研究所 ]。  
[http://ir.lib.ncu.edu.tw:88/thesis/view\\_etd.asp?URN=110454016](http://ir.lib.ncu.edu.tw:88/thesis/view_etd.asp?URN=110454016)
- 林芝儀(2020)。整合情感分析與文字探勘技術應用於新聞標題分析：以外匯市場預測為例 [ 碩士論文，國立臺北科技大學資訊與財金管理系 ]。  
<https://hdl.handle.net/11296/5ev38d>
- 孫嘉蔚 (2021)。運用機器學習模型分析影響公司風險的 ESG 因子：以台灣市場為例 [ 碩士論文，國立政治大學金融學系 ]。華藝線上圖書館。  
<https://doi.org/10.6814/NCCU202101589>
- 張育嘉(2022)非營利組織運用社群媒體對部門之影響——以彰化縣喜樂小兒麻痺關懷協會為例 [ 碩士論文，國立雲林科技大學企業管理系 ]。  
<https://hdl.handle.net/11296/7nndkb>
- 張蔭淥(2020)當即時通訊進入非營利組織溝通——以永和社區大學瓦礫溝願景社 LINE 群組為例 [ 碩士論文，國立政治大學傳播學院傳播碩士學位學程 ]。  
<https://hdl.handle.net/11296/794avt>

陳尊鈺(2011)非營利組織社群媒體行銷運用之研究—以 Facebook 為例〔碩士論文，中臺科技大學文教事業經營研究所〕。

<https://hdl.handle.net/11296/5nc7k5>

陳婉玲(2014)。國際經驗與人格特質對非營利組織捐款意願之影響〔碩士論文，長榮大學國際企業學系〕。華藝線上圖書館。

<https://doi.org/10.6833/CJCU.2014.00058>

楊雀(2023)。利己乎？利他乎？—人格如何透過動機影響親社會推薦行為的情感反應〔博士論文，淡江大學/商管學院/管理科學學系〕。華藝線上圖書館。

<https://doi.org/10.6846/TKU.2023.00070>

葉慈文(2023)。社群媒體遊戲對青年族群的毒品防制宣導成效影響研究：以 Instagram 為例〔碩士論文，國立雲林科技大學數位媒體設計系〕。

<https://hdl.handle.net/11296/9s6kg3>

蔡瑀(2023)。應用機器學習之留言分析系統設計：以 Youtube 教育影片為例〔碩士論文，國立臺北教育大學數位科技設計學系(含玩具與遊戲設計碩士班)〕。

<https://hdl.handle.net/11296/94cq5y>

鄭淑華、許哲瀚、龔建吉、張馨云、趙建蕾(2019)。結合科技接受模型及任務科技配適度探討護理人員使用社群媒體對工作績效之影響。醫務管理期刊，20(4)，267-287。[https://doi.org/10.6174/JHM.201912\\_20\(4\).267](https://doi.org/10.6174/JHM.201912_20(4).267)

盧憶慧、林清壽(2013)。中高齡者慈善捐款行為意向之研究—以南投縣為例。福祉科技與服務管理學刊，1(3)，35-50。

<https://doi.org/10.6283/JOCSSG.2013.1.3.35>

蕭丞勛(2020)。民眾捐款行為及捐款意向之研究—以嘉義市某身障安置機構為例〔碩士論文，國立暨南國際大學人文學院〕。華藝線上圖書館。

<https://doi.org/10.6837/ncnu202000341>

薛妙賢(2014)。關係行銷在非營利組織勸募關係運用之初探—以財團法人羅慧夫顏顏基金會為例〔碩士論文，朝陽科技大學人文暨社會學院〕。華藝線上圖書館。

<https://www.airitilibrary.com/Article/Detail?DocID=U0078-0905201416542664>