

經濟部 109 年度
《人工智慧產業關鍵技術拔尖計畫(3/4)》
合作研究計畫

《問答知識推論技術》

建議書徵求文件

財團法人資訊工業策進會

中華民國 109 年 00 月 00 日

109年度合作研究計畫建議書徵求文件

一、 簡介

問答知識推論旨在利用知識圖譜(Knowledge Graph)中的事實來回答自然語言問題。它可以幫助用戶更高效、更輕鬆地訪問知識圖譜中的重要知識。透過深度學習模型學習推理能力並且利用現有資源進行推理，學會推斷未知的答案。

為了推斷未知的答案，模型需要學習推理公式，例如：知識圖譜中包含內馬爾在巴塞羅那比賽以及巴塞羅那在西甲聯賽中的信息，因為”內馬爾”屬於”巴塞羅那”，而”巴塞羅那”又屬於”西甲聯賽”，那麼模型需要學習到內馬爾在西甲聯賽比賽的推理。透過學習知識圖譜中的推理路徑，模型能夠自動推斷一對具名實體之間的缺失鍵結，從而完善問答知識推論技術。

二、 計畫目標

知識圖譜推理是一種通過查詢應答任務中的推理路徑預測目標實體的有效且可解釋的方法。現階段僅能針對問句並往外進行one-hop查詢，針對取得之節點或關聯，決定反問策略。

相較於傳統模型，本計畫旨在透過少量的三元關係資料集，利用multi-hop對於問句的關聯所連接的兩個實體，快速了解知識圖譜中的關聯，預測它們之間的未知關係，並正確的推論回答。

三、 計畫範圍

運用深度學習模型建立可適用於中文知識圖譜的問答知識推論演算法：

- 需識別文件中提到的具名實體位置與類型並連結到知識圖譜
- 可從現有的知識圖譜關聯中，推論出未直接相連之三元關係

工作項目範圍包含：

- (1) 建構一特定領域知識圖譜，並研讀以深度學習相關問答知識推論演算法進行研究。
- (2) 擷取一目標文本，作為加入該領域知識圖譜之新增語料，並進行知識圖譜和具名實體之三元關係。
- (3) 處理目標文本，取得該文本中具名實體位置與類型和關聯標記。
- (4) 建立問答知識推論機制，從知識圖譜透過multi-hop選出可與目標問句中具名實體相連結的具名實體，與既有技術進行比較評估。
- (5) 建立問答知識推論模型，提供實體三元關係關聯預測機制。
- (6) 撰寫研究報告，含實驗評估模型有效性。

四、 預期成果(明確說明合作研究成果之產出)

- (1) 深度學習相關問答知識推論技術之演算法評估。
- (2) 特定領域知識圖譜及作為新增語料之目標問句及關聯。
- (3) 具名實體與關聯標記資料結構，及實體關聯資料集(含驗證與測試)。
- (4) 本案深度學習演算法與既有技術演算法之實驗比較。
- (5) 本案深度學習演算法程式雛型一份，包含實體三元關係關聯預測機制。
- (6) 研究報告一份。

※前述成果如有專利構想或專利申請產出時，需注意專利申請之新穎性(novelty)。因凡經公

開發表之研發成果，如擬申請專利，須於公開發表後6個月內完成，前述成果如是以論文方式公開發表，將無法取得大陸與歐盟等國之專利。(※文字請保留，此括號文字請於正式版時刪除)

五、執行方式(包括計畫時程、計畫分工方式)

查核點	時間	產出物	報告架構/大綱
M1	2020/06/15	期中報告	<ul style="list-style-type: none">■ 文獻探討■ 建立問答知識推論語料■ 問答知識推論演算法設計
M2	2020/11/30	研究報告 程式雛形	<ul style="list-style-type: none">■ 問答知識推論技術期末研究報告書 (期末報告需包含雛形軟體與既有技術之實驗比較)■ 雛形軟體成果與操作說明

六、計畫期程及預估計畫總經費

計畫執行區間：109年01月01日至109年12月31日

總經費：800,000元

七、驗收標準(含教育訓練)

- (1) 問答知識推論技術實作程式(包含雛形軟體與既有技術之實驗比較，包含實體三元關係關聯預測機制，特定領域知識圖譜及目標文本相關語料)。
- (2) 研究報告一份。

八、技術能力需求(請詳述所需要之技術能力或專長)

- 熟悉深度學習演算法設計研究問答知識推論技術，依據問句可透過知識圖譜之結構，進行至少2-hop之推論。
- 熟悉知識圖譜建構相關技術，包含實體關聯、維護補正。
- 具備機器學習和深度學習演算法設計經驗，與研究分析能力。
- 具備程式撰寫與軟體架構開發能力，實作演算法原型，並進行實驗量化驗證。

附件1：契約書格式

1-1：計畫書格式

1-2：經費動支報表

1-3：成果報告撰寫須知

1-4：報告格式

1-5：論文格式

1-6：保密聲明書

1-7：委託匯款同意書

