

腎臟超音波格式化專題

Renal Ultrasound Formatting Project

指導教授：蔣榮先

專題成員：蔡翔逸

開發工具：Chatgpt

測試環境：Chatgpt

一、簡介：

題目摘要：

以腎臟超音波為實務例子，將影像文字報告以 ChatGPT 的語意辨識與統整能力彙整成格式化資料，以用作後續可做為統計用的資料。

背景：

成大醫院已經成立三十多年，為臺南地區重要的醫學中心。其中累積的病患資料無論是數量還是質量都有非常好的規模，但因為原先儲存的資料型態都是文本資料，所以在進行統計研究時無法直接取用。如果能利用大型語言模型對於文句理解的能力將腎臟超音波報告轉換成可經一般程式讀取的檔案型態（Json 檔案），提高醫療資料的可讀性和可處理性，進而支援自動化的醫療資料分析和整合，為未來的醫療資訊系統提供強大的後盾。

介紹：

本次專題分為產生資料以及檢查錯誤兩個部分：

在生成資料的部分，首先建立一個合適的轉換後資料模板，包含以下項目：Medical history No.、Personal Information、Date of Examination、Kidney、Liver、Bladder、Ureters、Prostate、Other 以及 Impression。接著，通過完善且精確的提示詞（Prompt）來正確生成資料。發現初始提示詞的正確率不足 28%，因此將錯誤分為六大類，並通過 self-consistency 和 Direct prompting 進行錯誤改善，總共進行了五個版本的測試（每個版本測試 50 筆）。

在達到一定的資料轉換正確率之後，進行下一個步驟：檢查錯誤。這個步驟的主要目的是抓出剩餘的錯誤，以達到最大的資料正確率。檢查過程是透過分次分段的方式進行，並將欲檢查的錯誤同樣使用 self-consistency 和 Direct prompting 的方式加入提示詞（Prompt）中。總共檢查了 50 筆錯誤資料。

二、 測試結果：

在測試五版生成 prompt 後得到各自的正確率，version 1 正確率為 28%，version 2 正確率為 34%，version 3 正確率為 52%，version 4 正確率為 64%，version 5 正確率為 82% (最終版)，共提升了 54 個百分點。

版本	正確率
v1	28%
v2	34%
v3	52%
v4	64%
v5	82%

測試 50 筆錯誤資料的檢查之後，共可以檢查出 42 筆錯誤資料，可得檢查成功率達到 84%。

在得到 82% 的生成正確率以及 84% 的檢查正確率之後，把兩者結合。用 version 5 生成 prompt 生成資料之後，再透過檢查 prompt 將錯誤的資料檢查出來，我們將可以得到正確率超過 95% 的轉換後資料。