

國立台北大學資訊工程學系專題報告

Criminal Judgement Prediction Using Natural Language Processing and Artificial Intelligent Technology

應用自然語言處理與人工智慧技術於常用刑事類別判決預測

專題組員：葉宥麟 410685013、吳建宇 410685014

李侑程 410685037、黃品程 410685038

專題編號：PRJ-NTPUCSIE-108-010

執行時間：108 年 9 月 至 109 年 6 月

一、摘要

自然語言處理是現代技術最重要的組成部分之一，自然語言是指人們日常使用的言，是自然而然的隨著人類社會發展演變而來的語言，而不是人造的語言，它是人類學習生活的重要工具。概括說來，自然語言是指人類社會約定俗成的，區別於人工語言，如程序設計的語言。自然語言處理是計算機科學，信息工程和人工智慧的子領域，此領域探討如何處理及運用自然語言。自然語言處理包括多方面和步驟，基本有認知、理解、生成等部分。自然語言認知和理解是讓電腦把輸入的語言變成有意思的符號和關係，然後根據目的再進行處理，本專題希望透過電腦系統的判斷，模擬法院刑事案例判決，節省時間，節省人力，達到參考效果。

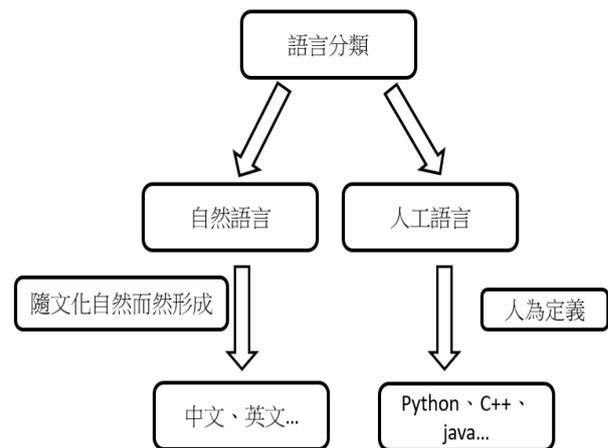


圖 1 語言分支架構

二、研究動機與研究問題

每個人一生當中，都會碰到需要法律諮詢的時候，一般來說，刑事案件可以分為自訴與公訴，但是不論自訴與公訴過程都非常繁瑣，且曠日廢時，從事件發生開始，一直到法院判決結果確定，時間更是無法預期，短則數周，長則數年，每天都有無數的案件需要等待法院進行審理，在有限的法律資源與現行的制度下，每個人所獲得法律幫助嚴重分布不

均，在資源允許的情況下，請一個專業的律師，當然是最理想的情況，然而，幫每一個需要法律協助的人都請一個律師是非常難實現的，因此如果可以做出一套刑事判決預測系統，可以提供給需要人，一個參考的預測結果，是一個非常節省資源的方法，且成本較經濟實惠，絕大部分的人都可以使用，主要的研究問題有以下：

1. 刑事種類選擇：

刑事種類非常多樣化，要是挑選適當的類別，主要以較貼近生活的類別為主。主要選擇青少年較常涉及的類別，分別為傷害、竊盜、詐欺、毒品。

2. 法律詞語的理解：

法律有自己的專業用語，與一般的自然語言不盡相同，理解上有一定落差。

3. 判決結果的判斷：

法院的判決結果非常多樣化，如何在複雜的法院裁判書，抽取結果

4. 預測的準確性：

成品的準確性預測非常重要，如何精準測試是一大問題。

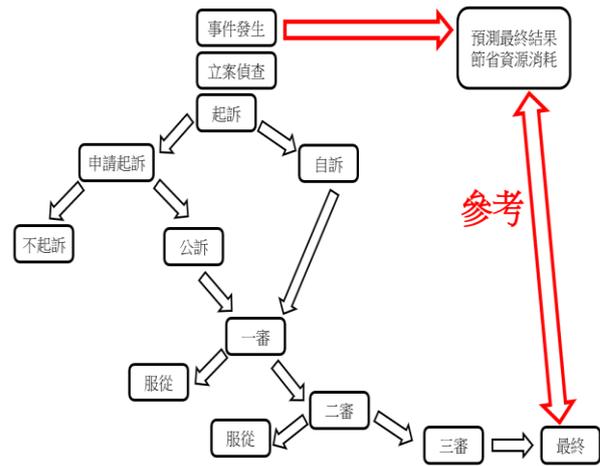


圖 2 節省流程架構

三、文獻回顧與探討

法律判決輔助程式最早在 1987 年由 Rissland 和 Ashley 開發(HYPO CBR)，直到今日許多各地也希望能夠使用人工智慧的方式協助判決，在中國每年都有比賽，而 2018 年則是以找出判決書的相關法條、刑期、罰金為主要目的[9]也是我們的企劃的主要目標。為了讓電腦理解人類的語言，需要將文本轉換成電腦看得懂的形式，因此 Word2Vec 便扮演著重要的角色，透過大量的文本資料訓練並使用數學向量來代表文意的相似度(cosine 值)，數值越高則代表詞之間的意義較相近。而除了基本的 Word2Vec 外，我們也嘗試了 Paragraph Vector 的方式[7]，訓練方式多了一個文檔的權重，可以使機器更加了解某詞在一篇文章中的詞序和詞意，在 input 方面也可以是變數，不同字數的文章或句子都可以輸入，無論

是文意的表達或是 input 的靈活度都表現得更好。而本計畫也嘗試理解許多種不同的模型來預測相關法條[8][10]，以標籤分類常見的刑事案件，由於刑事案件牽涉法條過於龐大再加上本企劃的規模，故以常見的幾項法條來作為預測目標。

四、研究方法及步驟

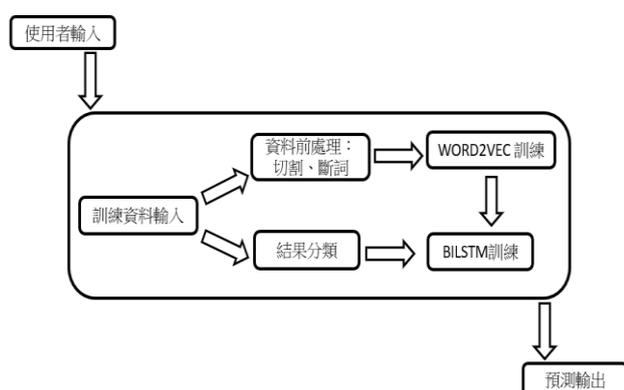


圖 3 預測步驟流程架構

1. 資料蒐集：

在用於整個研究過程的資料，會從司法院法學資料檢索系統取得，利用過去的裁判結果作為訓練和測試的資料。其中研究過程中裁判書的範圍，以各縣市的「地方法院」裁判的刑事案件為主，理由在於地方法院為第一審的事實審，判決主文比高等法院和最高法院更為明確，而刑事案件的裁判結果通常以刑期和罰金為主，因此相對於民事案件和行政案件更為清楚。

2. 裁判書文本錢處理：

從司法院法學資料檢索系統所提供的裁判書檔案形式，是無法直接拿來利用，因此必須要先進行文本的前處理轉換成電腦能夠處理的形式。解析完後的裁判書，存在特殊的紀錄格式，大致上可以分為六個部分，包括：裁判書裁判字號、與案件相關當事人資訊、主文、事實、理由和出席司法人員等，細節部分會因為記錄人員不同，而有些許不同。而本專題在研究過程所需的主要為主文、事實和理由這幾個部分，主文的部分能夠幫助我們或的裁判的結果，而事實和理由則是整個對於整個案件的陳述。在這幾個部分文本的差異不大，因此本計畫採用正則表達式，對文本進行分割，以文本中出現的特定字詞對其進行分割。

3. 斷詞斷句：

若是一般的英文的文本，其作為句子的基本單位是一個單字，因此以空白來做為分割的依據即可，但不同與英文的文本，中文的詞和詞之間並沒有一個明顯的分割字符，因此我們必須依賴額外的斷詞模型來處理，而現階段中文文本的斷詞模型，以 JIEBA 分詞和在 2019 年釋出的中央研究院的 CkipTagger 中文斷詞系統為主。而在這部分雖然 Jieba 相較於 CkipTagger 有比較方便的語料擴增自由

性，但 Jieba 分詞主要是以簡體字來進行建構的分析模型，且在各項衡量指標下都是以中研院的 CkipTagger 較為合適，因此在這個部分我們還是以訓練資料程度和較為精確的 CkipTagger 斷詞系統來做為本專題的斷詞模型。

Tool	(WS) prec	(WS) rec	(WS) f1	(POS) acc
CkipTagger	97.49%	97.17%	97.33%	94.59%
CKIPWS (classic)	95.85%	95.96%	95.91%	90.62%
Jieba-zh_TW	90.51%	89.10%	89.80%	--

圖 4 斷詞總類結果分析

4. 擷取判決結果：

與在分割裁判書的時候相同，因為主文只會包含有判決結果的資訊，因此我們透過判斷主文當中，是否出現「有期徒刑」、「拘役」和「罰金」等字詞來判斷被告是否有被判刑。

5. WORD2VEC：

在將字句斷詞之後，要將每一個字詞轉換為以數值向量表式的形式，而這部分 Google 2013 發表的 word2vec 模型，基於類神經模型來達到將字詞向量化的方法。不同於早期 One-hot encoding 的方法，One-hot encoding 分詞後的結果會是每一個單詞對應到一個位置，因此有 N 個單詞，則每個單詞就

會對應到一個 N 維向量，這樣結果容易會造成維度過大，而且詞和詞之間部會存在關聯性。而 word2vec 包含 Continuous Bag-of-Words 和 Skip-Gram 這兩種訓練模式(如下圖所示)，Continuous Bag-of-Words 利用選定的中間詞周圍的詞賴預測中間詞，至於 Skip-Gram 會根據選定的中間詞來預測其周圍的詞，用這兩種訓練模式的結果，可以有效的改善原先 One-hot encoding 的缺點，利用訓練後的結果，我們就能將字句轉成電腦能處理的形式。

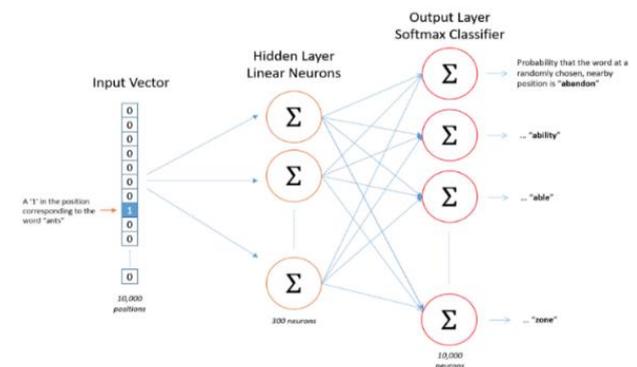


圖 5 word2vec 運作示意圖

6. BILSTM：

在 Natural Language Processing 的問題中，多數處理的是序列到序列的問題，在這樣的序列中字詞之間出現的順序就變得很重要，因為在中文的句子當中，有可能會有兩句話出現的字詞相同，但語意完全不同的情況。因此在 Natural Language Processing 的問題中，多半會採用 RNN 和 LSTM 這種遞歸神經

網路，在這類的神經網路模型中，會將現階段的輸出結果傳遞給下一個階段作為輸入使用，因此在進行下一個階段的處理時，會考慮到前面出現過的輸入。但若是在一般的 RNN 模型當中，若是序列的長度過大，會出現梯度爆炸的情形。因此在 LSTM 的模型當中，對 RNN 的模型進行改良，在神經元中增加了 Forget gate layer、Input gate layer 和 Output gate layer 三個層次。透過這三個層次能夠選擇經過的資料是否通過，只將有需要的資料留下來，不需要的捨棄掉。因此他可以比 RNN 更有效的運作在長序列的資料上。而若是一般的 LSTM 模型來說，只會考慮到句子之中已出現的部分，但很多的時候句子，後半部的訊息也是重要的，因此 BiLSTM 會結合前向的 LSTM 和後向的 LSTM，讓不管是句子中後半和前半的內容都考慮到。而透過這樣的將前處理後的結果，丟入 BiLSTM 模型，就能產生一個能預測判決的模型。

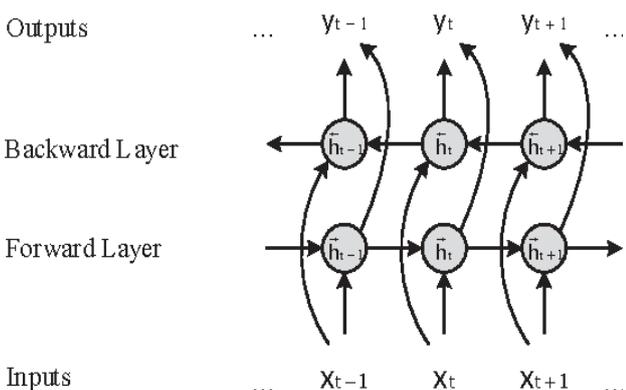


圖 6 BILSTM 運作示意圖

五、使用者輸入

讓使用者根據現實事件發生的經過與細節，輸入系統，輸入的方式為全手動輸入，為了更加貼近一般民眾生活，輸入時會有相關範本參考，且輸入內容中不需要含有相關的法律知識，與法律相關用語，讓使用者輸入，以簡單的敘述方式即可。

六、輸出結果

依據使用者的輸入，輸出三個預測項目

1 所涉及的刑事種類

依據使用者的輸入事件，輸出該事件涉及的刑事種類。

2. 可能的有罪無罪結果

依據使用者的輸入事件，輸出事件有罪無罪的機率。

3. 可能的刑期

依據使用者的輸入事件，輸出事件如果被判刑的刑期長短。

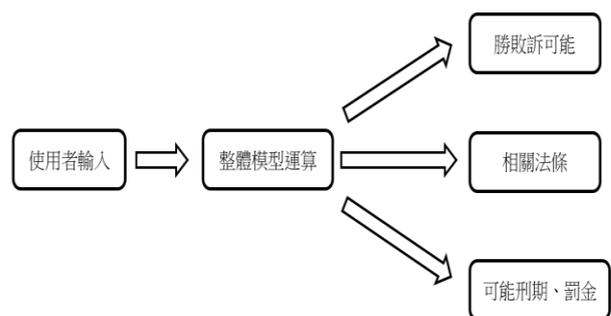


圖 7 預期輸出流程項目

七、預測結果評估

同樣以司法院所公布的裁判書中，以隨機的方式選取僅含事實(不會參雜法律專業知識)的裁判案例。

1. 此為分類模型結果(為一個大模型)

類別	傷害	竊盜	詐欺	毒品	總計
訓練數	29900	29900	29900	29900	119600
測試數	300	300	300	300	1200
正確數	297	391	297	297	1176
錯誤數	6	9	6	3	24
正確率	0.98	0.97	0.98	0.99	0.98

圖 8 分類模型結果

2. 此為預測有罪無罪模型結果(每個類別一個模型，一共四個)

類別	傷害	竊盜	詐欺	毒品	總計
訓練數	29900	29900	29900	29900	119600
測試數	300	300	300	300	1200
正確數	291	281	291	294	1158
錯誤數	9	18	9	6	42
正確率	0.97	0.94	0.97	0.98	0.965

圖 9 預測有罪無罪模型結果

3. 此為預測刑期模型結果(每個類別一個模型，一共四個)，表格括號中為預測正確數量，每一個類別測試三百篇

類別	傷害(天)	竊盜(天)	詐欺(天)	毒品(年)
1	0~10 (54)	0~30 (57)	0~90 (18)	0~1 (66)
2	11~30 (60)	31~90 (66)	91~180 (30)	1~2 (42)
3	31~60 (33)	91~150 (42)	181~270 (36)	2~3 (27)
4	61~120 (21)	151~210 (27)	270~360 (69)	3~4 (33)
5	121~180 (42)	211~270 (33)	360~450 (57)	4~5 (27)
6	>181 (33)	>271 (15)	>451 (18)	>5 (48)
準確率	81%	82%	76%	81%

圖 10 預測刑期模型結果

八、參考文獻

[1] Wang, An-Ding .(2016, 8). Text Mining of Court Judgments and Regression Model for Judicial Sentence: An Example from Judgments on Narcotics Crimes

[2] Chih-Wen Su .(2017, 7). 2000~2012 Big Data Analysis on the Judgment of Criminal Medical Malpractice in District Courts in Taiwan from 2000 to 2012 - Predicting the Judgment with Machine Learning

[3] LIN, SHIAN-CHING .(2018, 7). The Design and Implementation of a Sentencing Prediction

[4] CHI, HSING-CHEN .(2018, 10). A Machine Learning Based Prediction System of Medical Laws Judgment Using Statute-Classified Decision Tree with Text Similarity

[5] Shaio Kuan Lu .(2019) Nurse and Patient Dispute Litigation in Nursing Practice: Analysis of District Courts'

Criminal Judgments

[6] Chih-ChiaChang. (2019). Empirical Study on Compensation for Damage of Trademark Infringement: Extracting Data from Court Verdicts with Python

[7] Distributed Representations of Words and Phrases and their Compositionality. Retrieved from <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf> (2020,2)

[8] A Recurrent Attention Network for Judgment Prediction. Retrieved from https://www.researchgate.net/publication/335699354_A_Recurrent_Attention_Network_for_Judgment_Prediction (2020,2)

[9] Overview of CAIL2018: Legal Judgment Prediction Competition. Retrieved from <https://arxiv.org/abs/1810.05851> (2020,2)

[10] Bidirectional LSTM-CRF Models for Sequence Tagging .Retrieved from <https://arxiv.org/abs/1508.01991> (2020,2)