

交通事故法規分析

指導教授: 鮑興國 組員: 柯映竹

利用交通事故案件之文字描述，分析出案件適用之法規

1 Introduction

此專題目標為預測交通事故肇事適用之法律，根據不同縣市之交通事故文字紀錄，利用機器學習的文字探勘及其他相關工具預測案件相關法規。

2 Methodology

方法一: 駕駛行為直接預測法規

使用分類演算法，直接進行預測

步驟:

- 對駕駛行為做encoding
- 使用MODEL 進行訓練及預測法規

問題:

- 每一案件有多條相關法律，但是未找到多標籤的模型
- 直接使用資料，沒有更深入的觀察

方法二: 駕駛行為-標籤-法規

將駕駛行為貼上相關之標籤，每一標籤會對應到相關之法條

步驟:

使用TOPIC MODELING-LDA分群演算法

- 分群法律
- 分群駕駛行為
- 觀察兩個分群結果，重新分類，以及連結標籤

問題:

兩者分群結果差異大，無法以共同標籤連結起來

方法三: 分群駕駛行為

分群駕駛行為，再參考同群之法律

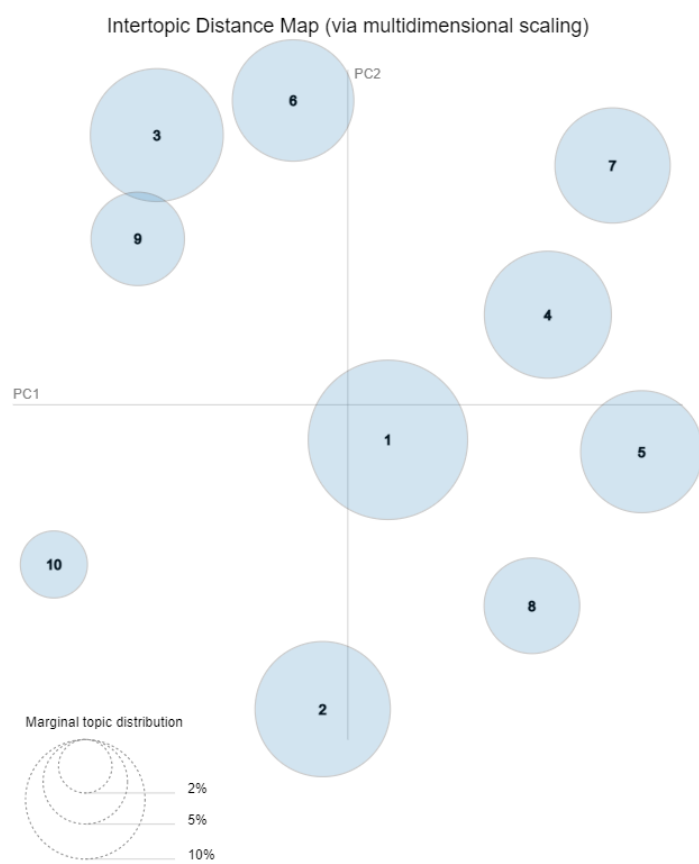
步驟:

使用TOPIC MODELING-LDA分群演算法

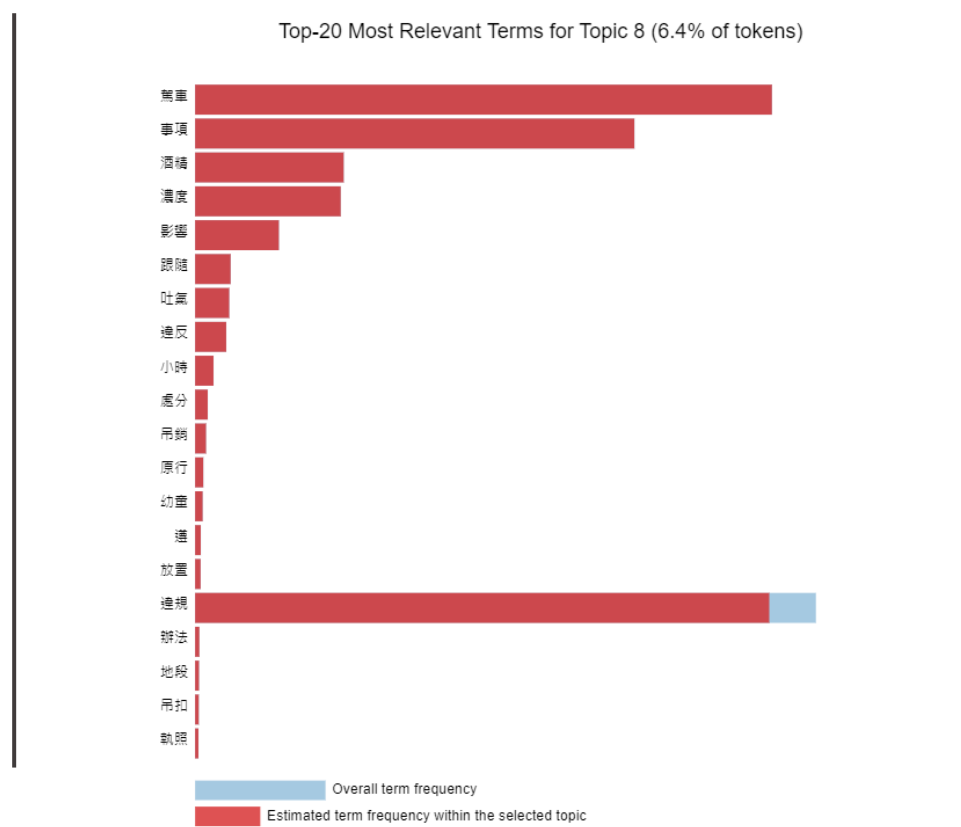
- 分群駕駛行為
- 參考同一群使用之法律

最後使用方法3。

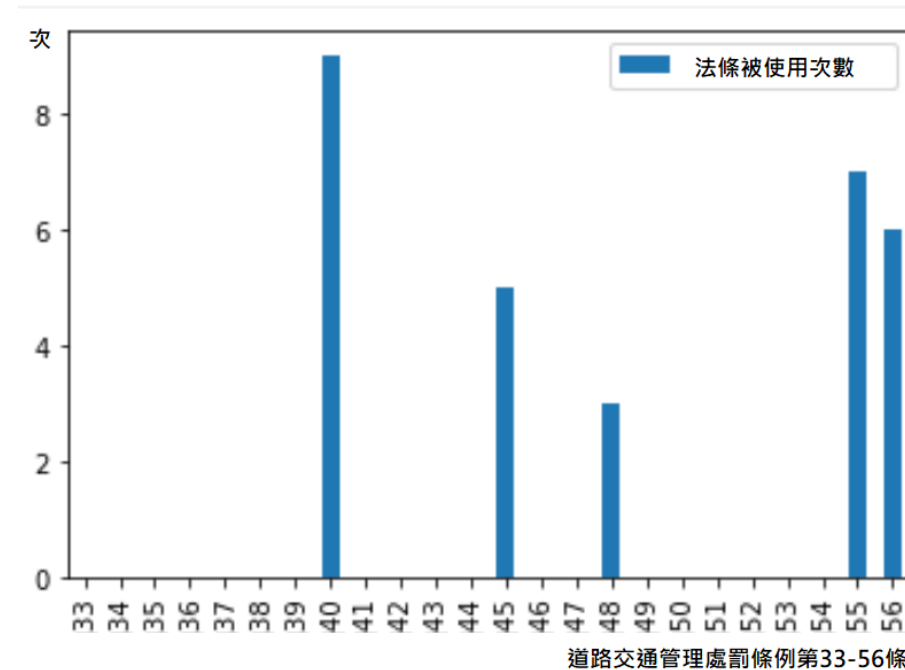
5 Analysis



分群結果圖示



每一群的關鍵字(圖為TOPIC8)



與 TOPIC 8 有關之法律使用次數

6 Conclusion

利用Topic Modeling 演算法，搭配上CKIP斷詞工具，可以預測案件適用法規，並且解決分類演算法中只能使用一對一預測模型的問題。其中關鍵詞的採用依據若使用法律斷詞及駕駛行為斷詞的交集詞彙，其分群的結果較有應用性及研究性。

未來可以進行的優化:

- 擴大案件使用之法律:

目前僅針對交通事故管理處罰條例第33-56條進行預測，未來可以擴增至道路交通安全規則及道路交通號誌標線號誌設置規則。

- 參考其他欄位，如肇事因素、車種等等。
- 使用多層分群，觀察是否有更準確的預測。

另外，在分群過程中，發現同群之中，有看起來有隱藏之關聯的主題關鍵字。例如:

- 逆向、超車、計程車在同一群當中，是否有隱藏的相關性？
 - 闖紅燈與左轉彎在同一群，左轉彎占闖紅燈案件的比例是否最大？
 - 甚至有些路名在特定群體的出現頻率高，是否與其路段設置有關？
- 這些分群的結果都很值得繼續深入套討。