

作品集

作品一：機器學習/數據處理(以探討學生學習情形為例)

作品二：智法助手-法律RAG系統問答機器人



作品一

機器學習/數據處理(以探討學生學習情形為例)



目錄

1. 背景與動機
2. 痛點與目的
3. 數據資料說明
4. 模型介紹
5. 效能分析
6. 困難與改善方向
7. 結論
8. 參考資料

1.背景與動機

背景動機：

少子化、課綱改革、學生閱讀能力下降...等問題，補習班面臨招生問題。

思考方向：

藉由觀察歷次成績、在班時間、續班率、試聽報名情形等因素，與學生成績走向相關性，探討學生成績受各項因素影響的程度。

2.痛點與目的

痛點：

學生對於學校教學吸收不完整，家長看到其他孩子到補習班獲得良好成績，對於補習班產生需求

目的：

<輸入>擷取補習班國二生一段期間的成績紀錄

<模型>採用線性回歸(Linear Regression)當訓練模型

<輸出>回歸圖形、MSE值、R-Squared值

預期效益：

找出小考成績之間的相關性，預測學生未來成績，方便與家長溝通可能出現的狀況，進一步提高與家長的連結性，增加續班率，提高公司收益

3.數據資料說明

資料蒐集來源：

業界老師真實考試成績登錄

資料介紹：

擷取國二生一段時期的成績紀錄，包含小考與三次段考

資料型態：

數值(Numeric Types)，整數(int)為主

資料前處理-1

學生請假空白部分，以整體平均值填入，以免效能降低

修正前

62	40	70
100	76	85
100	假	80
94	73	60
74	28	57
56	20	54
91	76	77
63	68	54
96	65	假
..

修正後

62	40	70
100	76	85
100	77	80
94	73	60
74	28	57
56	20	54
91	76	77
63	68	54
96	65	86
..

資料前處理-2

3次RS值代表3次段考，我們將這3次的數值移至最後方便一起討論

7test_20	1026test_21	1107test_22	1012RS_1	0314RS_2	1003RS_3
81	68	88	60	70	77
52	72	81	88	74	66
77	76	92	100	88	81
88	92	95	100	96	85
73	80	82	78	72	58
88	92	92	100	88	90
88	92	92	92	92	90
40	62	60	62	88	70
76	84	93	65	88	85
77	80	88	77	92	80
73	80	81	71	85	60

由於studentID那一行沒有任何相關性,所以把這行drop,僅將成績當訓練集/測試集

3	1104	100
4	1105	63
5	1106	87
6	1107	78
7	1108	61
8	1109	40
9	1110	80
10	1111	75
11	1112	55

<處理後資料>
橫軸：考試資料
縱軸：學生ID

studentID	0830test_1	0920test_2	0927test_3	1004test_4	1101test_5	1115test_6	1120test_7	1122test_8	1201test_9	...	0515test_16	0621test_17	0808test_18	0829test_19	0927test_20	1026test_21	1107test_22	1012RS_1	0314RS_2	1012RS_3
1101	25	33	43	35	61	28.0	38	68	62	...	76	55	57	76	61	68	88	50	70	70
1102	56	35	68	66	56	59.0	66	77	62	...	80	68	67	80	52	72	81	86	74	74
1103	89	75	85	87	78	57.0	68	98	78	...	92	88	84	79	77	76	92	100	88	88
1104	100	85	92	96	74	64.0	84	100	82	...	100	88	92	92	88	92	95	100	96	96
1105	63	45	48	71	52	48.0	52	63	63	...	64	60	42	78	73	80	82	76	72	72
1106	87	45	85	94	88	72.5	70	96	91	...	88	91	94	100	88	92	92	100	88	88
1107	78	75	77	76	74	93.5	75	95	82	...	88	80	93	86	88	92	92	92	92	92
1108	61	63	55	71	67	33.0	49	64	68	...	96	88	62	62	40	62	60	62	86	86
1109	40	46	81	55	53	36.0	43	74	74	...	92	66	84	100	76	84	93	65	86	86
1110	80	80	81	91	82	85.0	86	88	91	...	88	66	81	100	77	80	88	77	92	92
1111	75	85	77	90	60	61.0	71	63	55	...	96	63	74	94	73	80	81	71	85	85
1112	55	51	32	66	25	55.0	51	42	43	...	36	50	31	74	28	56	56	61	57	57
1113	35	65	44	42	65	66.0	30	55	59	...	64	36	20	56	20	56	63	23	78	78
1114	81	86	77	91	86	83.0	91	64	75	...	84	88	80	91	76	92	91	92	88	88
1115	56	66	61	57	81	52.0	53	66	82	...	51	68	60	63	68	64	66	40	76	76
1116	86	87	90	84	66	87.0	100	76	94	...	79	96	81	96	65	84	89	92	70	70
1117	40	42	48	60	61	63.0	67	52	75	...	60	65	60	61	66	72	72	55	83	83
1118	44	43	61	12	29	30.0	55	28	34	...	32	33	33	25	48	16	33	22	50	50
1119	42	33	61	55	43	38.0	50	52	56	...	50	45	46	38	60	56	71	53	45	45

```
df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
0830test_1	19.0	62.789474	21.583376	25.0	43.0	61.0	80.50	100.0
0920test_2	19.0	60.000000	19.717449	33.0	44.0	63.0	77.50	87.0
0927test_3	19.0	66.631579	17.945010	32.0	51.5	68.0	81.00	92.0
1004test_4	19.0	68.368421	22.522114	12.0	56.0	71.0	88.50	96.0
1101test_5	19.0	63.315789	17.632207	25.0	55.5	65.0	76.00	88.0
1115test_6	19.0	58.473684	19.909848	28.0	43.0	59.0	69.25	93.5
1120test_7	19.0	63.105263	18.764312	30.0	50.5	66.0	73.00	100.0
1122test_8	19.0	69.526316	19.743827	28.0	59.0	66.0	82.50	100.0
1201test_9	19.0	69.789474	16.273970	34.0	60.5	74.0	82.00	94.0
0110test_10	19.0	67.684211	19.293213	35.0	51.5	71.0	82.00	100.0
0118test_11	19.0	67.052632	15.865664	31.0	60.5	69.0	76.50	92.0
0221test_12	19.0	70.236842	19.498126	28.0	55.5	72.0	84.50	100.0
0321test_13	19.0	77.052632	15.046128	43.0	69.0	84.0	88.00	96.0
0418test_14	19.0	66.736842	18.177648	30.0	55.0	61.0	80.50	96.0
0509test_15	19.0	70.473684	14.751152	44.0	59.5	71.0	84.00	89.0
0515test_16	19.0	74.526316	20.748142	32.0	62.0	80.0	90.00	100.0
0621test_17	19.0	68.105263	18.905892	33.0	57.5	66.0	88.00	96.0
0808test_18	19.0	65.315789	22.501982	20.0	51.5	67.0	82.50	94.0
0829test_19	19.0	76.368421	21.211136	25.0	62.5	79.0	93.00	100.0
0927test_20	19.0	65.473684	19.732569	20.0	56.0	73.0	77.00	88.0
1026test_21	19.0	72.315789	18.526415	16.0	63.0	76.0	84.00	92.0
1107test_22	19.0	78.157895	16.520234	33.0	68.5	82.0	91.50	95.0
1012RS_1	19.0	69.315789	24.700546	22.0	54.0	71.0	92.00	100.0
0314RS_2	19.0	77.684211	14.364511	45.0	71.0	83.0	88.00	96.0
1003RS_3	19.0	70.789474	14.581491	42.0	57.5	77.0	83.00	90.0

➤ 顯示19位學生在每場考試中整體表現，如最小值、最大值、平均值、標準差的結果，便於更快速瞭解數據集分佈的中心趨勢、分散程度和形狀。

4. 模型介紹-1

- 考量到家長期待成績表現上通常對於月考成績會特別關注、且月考成績是平時小考累積的展現，所以最後我們採用一學期的23次考試當訓練集，最後一次月考成績當測試集。

```
[ ] X = df.iloc[:, 1:24].values.tolist()
    y = df.iloc[:, 24].values.tolist()
```

```
[ ] X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=10)
```

```
[ ] linear_regressor = LinearRegression()
    linear_regressor.fit(X_train, y_train)
    y_pred_lreg = linear_regressor.predict(X_test)
```

4.模型介紹-2

- 尋找**最佳擬合線**。

使用線性迴歸時，我們的主要目標是找到**最佳擬合線**，這意味著預測值和實際值之間的誤差應該最小化，使最佳擬合線的誤差來到最小。

- 線性迴歸演算法顯示應變數 (y) 和一個或多個自變數 (x) 之間的線性關係，因此稱為線性迴歸。其特性可以發現應變數的值是如何根據自變數的值產生變化的，此特性符合小考與月考的關係，所以後來我們選用這組模型下去運用。

- 公式： $Y = a_0 + a_1 X + \varepsilon$

Y= 因變數 (目標變數)

X= 自變數 (預測變數)

a_0 = 線的截距 (提供額外的自由度)

a_1 = 線性迴歸係數 (每個輸入值的比例因子)。

ε = 隨機誤差

x 和 y 變數的值是線性迴歸模型表示的訓練資料集。

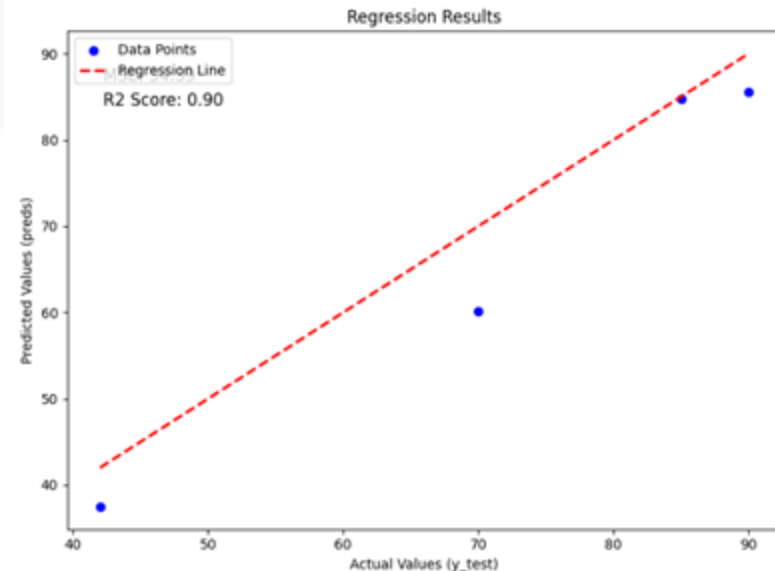
4.模型介紹-3

```
plt.figure(figsize=(8, 6))
plt.scatter(y_test, y_pred_lreg, c='blue', label='Data Points')
plt.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], color='red', linestyle='--', lw=2, label='Regression Line')
plt.xlabel('Actual Values (y_test)')
plt.ylabel('Predicted Values (preds)')
plt.title('Regression Results')
plt.legend()

plt.text(0.05, 0.9, f'MSE: {mse:.2f}', transform=plt.gca().transAxes, fontsize=12)
plt.text(0.05, 0.85, f'R2 Score: {r2:.2f}', transform=plt.gca().transAxes, fontsize=12)

plt.tight_layout()
plt.show()
```

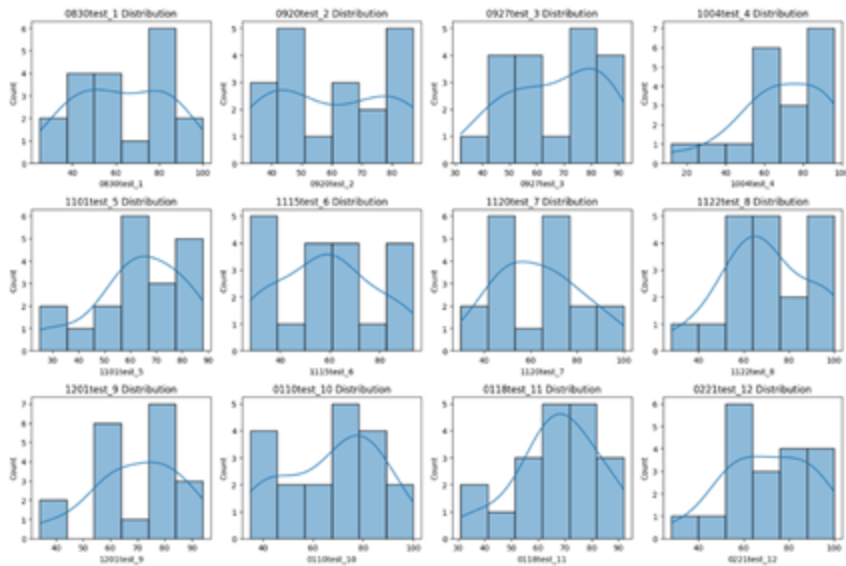
- 顯示回歸分析結果的圖形，包括實際值與預測值的散點圖、理想回歸線、以及模型性能的相關指標（MSE 和 R-squared 值）。
- 藍色散點:展示實際值 (y_{test}) 和預測值 (y_{preds})。
- 紅色虛線:代表理想的回歸線，即預測值和實際值完全一致的情況。



```
fig, axs = plt.subplots(nrows=3, ncols=4, figsize=(15, 10))
plt.subplots_adjust(hspace=1)

# Iterating through columns and create histograms
for col, ax in zip(df.columns, axs.ravel()):
    hisplot = sns.histplot(data=df, x=col, ax=ax, kde=True)    # Add KDE (Kernel Density Estimate) plot
    hisplot.set_title(f'{col} Distribution')

# Adding some additional adjustments
plt.tight_layout()    # Automatically adjust subplot parameters for better layout
plt.show()
```



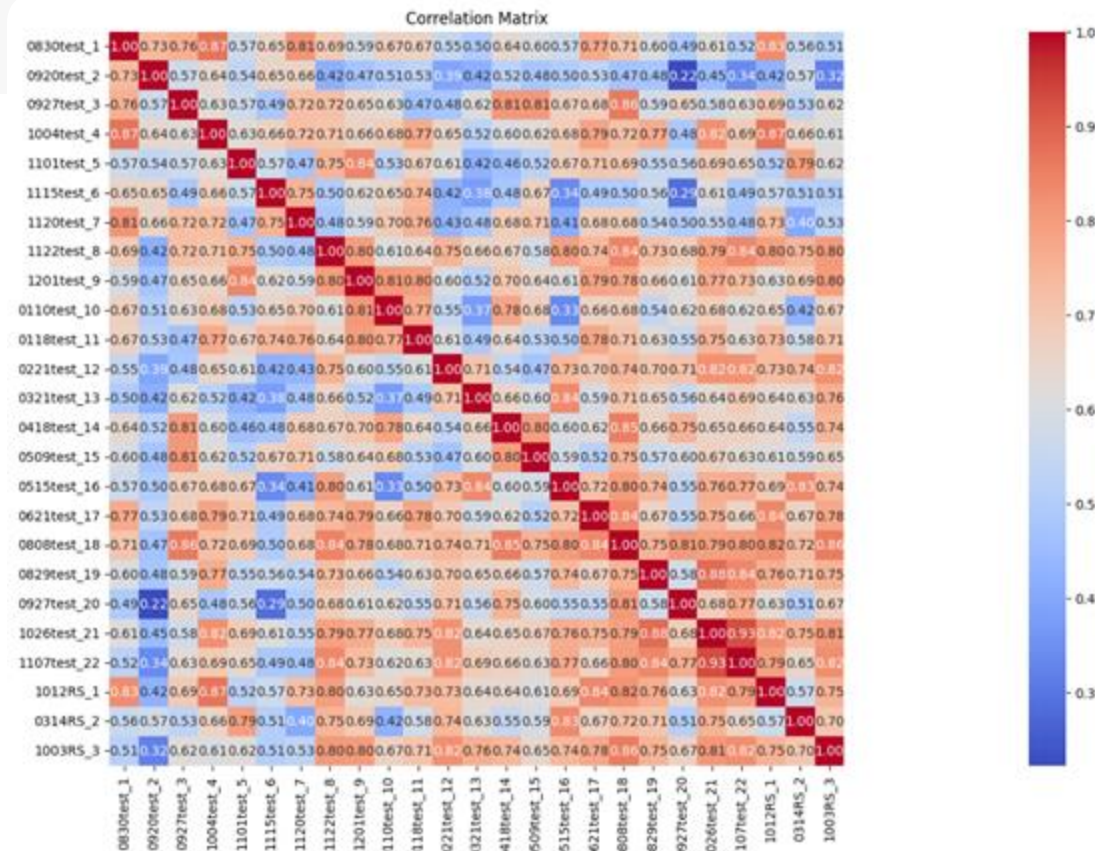
4.模型介紹-4

- 針對學生歷次考試成績分布，創建多個直方圖，並結合核密度估計（KDE）提供更豐富的數據視覺化。


```
corr_matrix = df.corr()
plt.figure(figsize=(50, 10))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f", square=True)
plt.title('Correlation Matrix')
plt.show()
```

4.模型介紹-5

- 繪製旋轉矩陣，將學生歷次考試成績相互比對，使成績表現的相互關係視覺化。



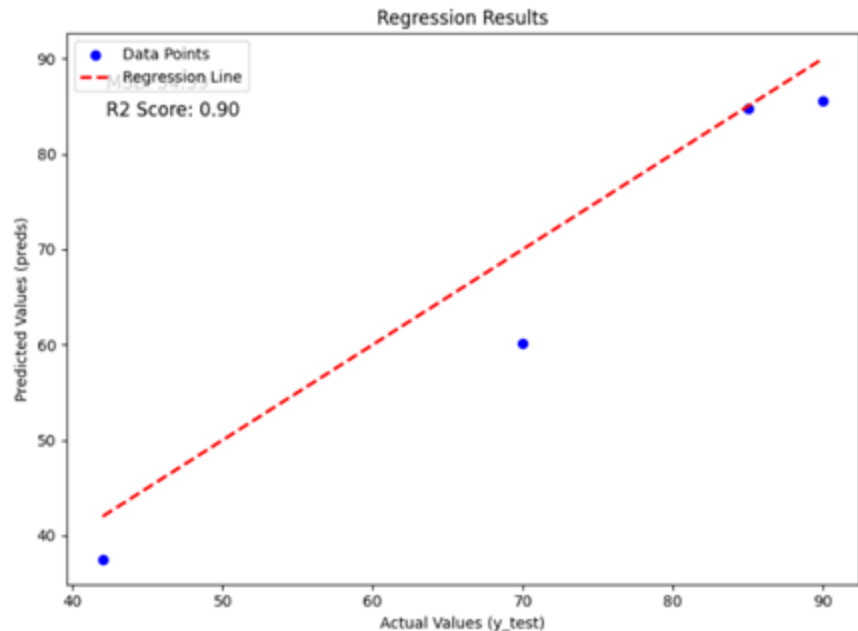
5.效能分析-真實數據結果

```
mse = mean_squared_error(y_test, y_pred_lreg)
r2 = r2_score(y_test, y_pred_lreg)
print(f'Mean Squared Error: {mse:.2f}')
print(f'R-squared Score: {r2:.2f}')
```

Mean Squared Error: 34.39

R-squared Score: 0.90

- 均方誤差（Mean Squared Error, MSE）和 R 平方（R-squared）值兩個指標來量化模型的性能。

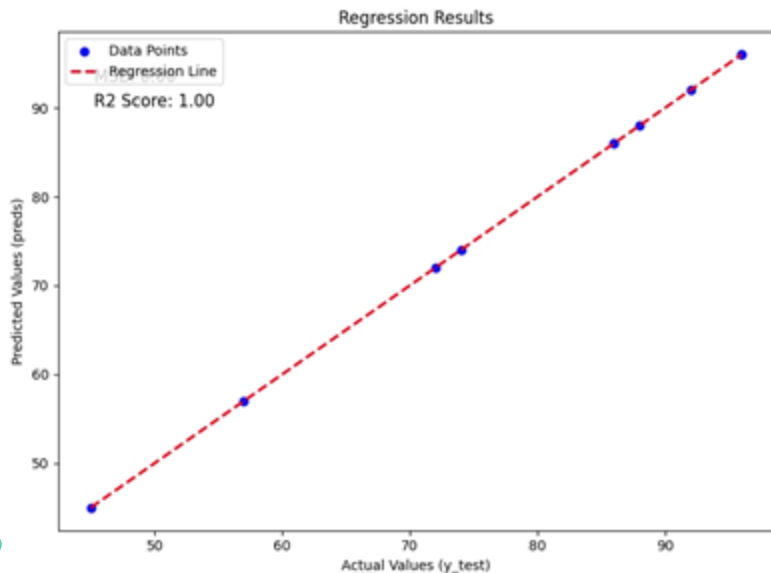


- 有4個點落在回歸線附近
- 平均誤差MSE值 = 34.39
- R-squared = 0.9

5.效能分析-生成數據結果

```
[ ] mse = mean_squared_error(y_test, y_pred_lreg)
    r2 = r2_score(y_test, y_pred_lreg)
    print(f'Mean Squared Error: {mse:.2f}')
    print(f'R-squared Score: {r2:.2f}')
```

Mean Squared Error: 0.00
R-squared Score: 1.00



➤ 思考點：為何真實數據出來效果不佳？

➤ 推測點：

(1)數據量不夠

(2)成績之間關連性並沒有想像的那麼高

(3)數據起伏太大

➤ 測試結果：

將原本數據複製多次之後，重新跑模型，得到
如圖藍色散點高度集中的完美狀態

平均誤差MSE值 = 0

R-squared = 1.0

所以應是[數據量不夠]

6. 困難與改善方向

遭遇困難

- 短時間內資料蒐集不易，且許多資料涉及個資問題。
學生成績不完全由成績起伏決定，上課方式、上課時長、學生家庭狀況...等因素都有影響，諸多影響因子無法量化，且變化快速。
- 複雜模型編碼建置及維護不易。

改善方向

- 可能加入更多影響因子:學生進班時間長短*、學生心理及情緒變化、家長教育水平及住家與學校及補習班距離遠近等。
- 擴充數據資料庫:用google表單的方式對學生進行行為和情緒分析或透過收集更大量且時間區段更長的成績表現，投入模型進行訓練。

7. 結論

- 成績與成績之間有存在相關性，但須輔以大量的數據，可使模型預測的結果更加精確。
- 家長最在乎的月考成績確實跟平時小考存在相關性。
- 以當前**少子化**和**小班教學**環境分析，海量的數據較難在真實狀況下實現，所以若要穩固**續班率**，我們討論出以下幾點可能：
 - (1) 學生家庭經濟狀況
 - (2) 學生個人心理狀態
 - (3) 回家作業難易度及負荷狀況
 - (4) 同儕效應
- 若能將以上幾點量化並加入模型，預期應該能取得除了成績以外，不一樣的相關性，進一步幫助預測**續班**的可能性。

9. 參考資料

- <https://www.kaggle.com/code/rustamovamalak/student-score-regression/notebook>
- <https://www.kaggle.com/code/spscientist/student-performance-in-exams>
- <https://www.javatpoint.com/linear-regression-in-machine-learning>
- <https://ithelp.ithome.com.tw/m/articles/10218158>
- <https://medium.com/ai反斗城/evaluation-metrics-迴歸模型-807c04871bb4>
- <https://bc165870081.medium.com/regression回歸問題的評估指標-4fb9b7c9a993>



作品二

智法助手：
法律RAG系統問答機器人

目錄

1. 簡介
1.1 背景、痛點、動機、目的
2. 系統架構、技術
2.1 數據資料說明
2.1.1 資料收集、來源與介紹
2.1.2 資料型態、前處理
2.2 模型介紹
2.2.1 應用模型之介紹及優缺點
2.2.2 選擇此模型之原因、優劣勢
3. 效能評估
4. 結論
4.1 Review技術、說明此專題的創新、貢獻
4.1.1 Review技術
4.1.1 創新與專題貢獻
4.1.1 專案價值與應用前景
5. 參考文獻
6. 補充、實驗資料

1.1 產業痛點分析

資訊分散

各項法規相互依賴關係繁多且分散，使用者難以及時獲取相關資訊。

專業門檻高

法條內容較為艱澀，對非相關背景人士而言理解困難，常導致誤解或未能行使合法權益。

資源不足

法律諮詢和法學教育成本高昂，對於企業和個體戶負擔較大。

法律風險

企業和個體戶因不了解法律規範，容易錯過第一時間尋求相關協助。

1.1 現有解決方案

1

專業律師或顧問服務

雖精準，但昂貴且無法即時應對大量需求。

2

政府網站或手冊

資訊呈現與內容制式，難以滿足個性化需求。

3

自我學習與查找相關資料

效率低、操作門檻高，且不易判別查詢結果是否適宜。

1.1 專案動機

1

降低法律諮詢門檻

提供即時解答，讓企業與一般使用者快速掌握法規內容，避免不必要的法律風險或爭議。

2

提升用戶法律意識

透過自動化工具普及法律規範意識，讓企業與一般使用者在第一時間快速利用人工智慧工具。

3

實現數位轉型

利用人工智慧優化勞工法規和職安相關的教育與諮詢流程，推動法律服務的智能化。

1.1 系統目的與功能

● 輸入

使用者提問（自然語言輸入）

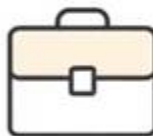
使用者所處的地區或行業（選填）

相關數據（法規條文、常見問題資料庫、指引等）

● 輸出

精準解答：以簡單易懂的語言提供法規相關說明

行動建議：如糾紛處理步驟、安全隱患應對方法



2.1 專案流程



- 資料來源：法務部「全國法規資料庫」的熱門法規
 - 資料蒐集方式：Python 網路爬蟲技術 + 迴圈抓取熱門法規 (共20部)
 - 將法規文本資料轉換為詞向量，並輸出成 JSON 檔案保存
- | | | |
|---------|---|--|
| 詞向量資料庫1 | → | chunk：一條法條；overlap：無
使用模型：sentence-transformers/all-MiniLM-L6-v2 |
| 詞向量資料庫2 | → | chunk：一條法條；overlap：無
使用模型：lianghsun/Llama-3.2-Taiwan-Legal-3B-Instruct |
| 詞向量資料庫3 | → | chunk：三條法條；overlap：一條法條
使用模型：lianghsun/Llama-3.2-Taiwan-Legal-3B-Instruct |
- 利用 Taiwan-Legal-3B-Instruct 模型 + 檢索增強生成 (Retrieval Augmented Generation, RAG) 技術
 - 從「律師事務所法律諮詢問答」及「全國法規資料庫常見問答」等選取10個問答題，進行應用模型效能評估 (人工檢核)
 - Line 聊天機器人

2.1.1 資料蒐集、來源與介紹

- 本專題使用20部台灣常見的法典作為RAG技術的外部資料。
- 法條內容則利用網路爬蟲技術擷取全國法規資料庫中的熱門法規。
- 為了獲取法條內容時保留其法典來源和法條資訊，我們將法規名稱、法條分類和法條條號保留於法條內容之前。

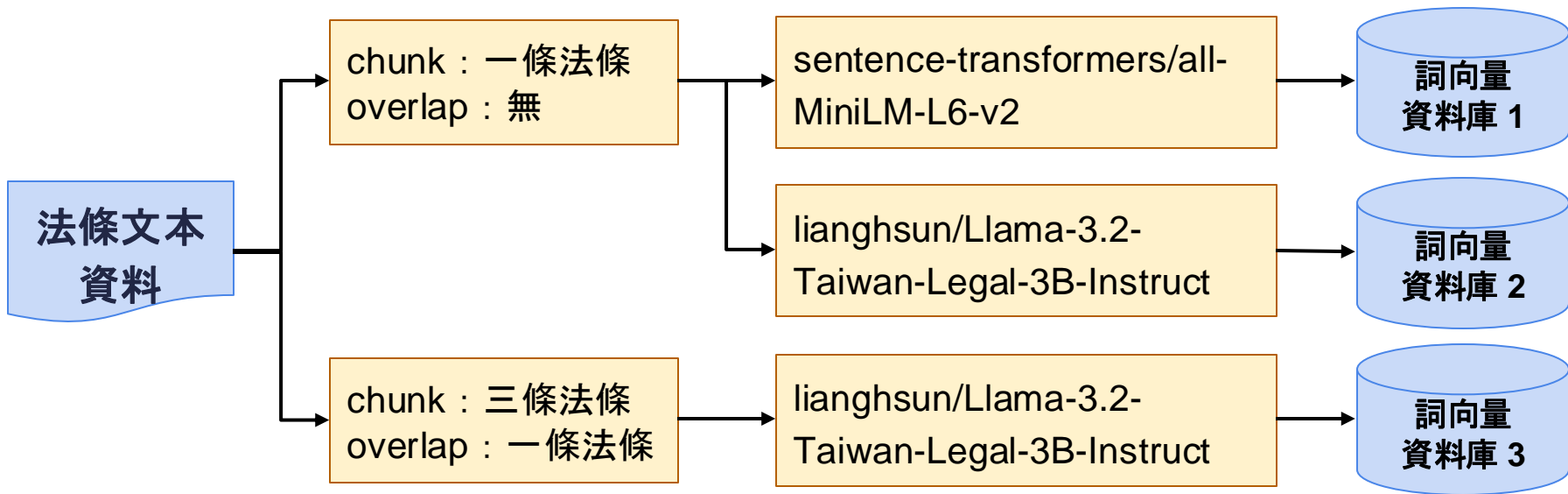
民法 第一編總則 第二章人 第一節自然人 第23條 因特定行為選定居所者，關於其行為，視為住所。

民法 第一編總則 第二章人 第一節自然人 第24條 依一定事實，足認以廢止之意思離去其住所者，即為廢止其住所。

民法 第一編總則 第二章人 第二節法人 第一款通則 第25條 法人非依本法或其他法律之規定，不得成立。

2.1.2 資料型態、前處理

利用爬蟲技術抓取的法條資料為文字型態，這些法條文本接著經過兩種不同切割方法及 *sentence-transformers/all-MiniLM-L6-v2* 和 *Llama-3.2-Taiwan-Legal-3B-Instruct* 兩個大型語言模型的詞嵌入 (word embedding) 處理後，將文字資料轉換為詞向量資料，並輸出成JSON檔案保存。



2.2.1 應用模型介紹

Taiwan-Legal-3B-Instruct模型是基於Llama-3.2-Taiwan-3B模型並利用台灣的法律條文與相關判決書資料集進行微調建構而成的模型，其發佈在Hugging Face平台上是開源的大型語言模型。Taiwan-Legal-3B-Instruct模型微調使用的資料集內容涵蓋法律條文結構、判決書格式，法庭上常見的法律語言與術語，同時包含部分法律資料科學任務的應用。

優點

相較於Meta公司開源的Llama-3.2-Instruct模型，經過繁體中文、台灣法條和判決微調後的模型更適合回答台灣法律相關問題，且對於繁體中文的使用更為流暢。

缺點

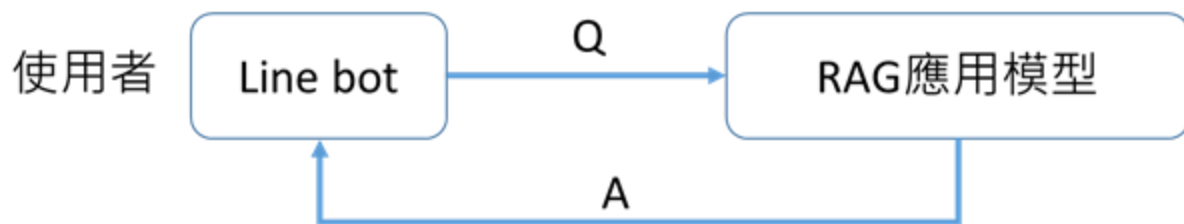
- 1、可能生成虛構不實的內容，使用者需謹慎審視模型回答的內容。
- 2、對於較複雜或語言模稜兩可的法條內容或案例事實，模型可能無法生成精確地回答。

2.2.2 選擇模型原因

- Taiwan-Legal-3B-Instruct模型經過台灣的法律條文與相關判決書資料集微調，更加適合回答台灣法律相關的問題。
- Llama系列以及各類型基於Llama的微調模型可以在Hugging Face上取得模型下載和使用的權限，方便在本地端開發和大量測試模型問答使用。
- 三十億參數規模 (3B) 的大型語言模型適合在本地端開發和使用，對於硬體設備的需求門檻較低，而更大參數規模的大型語言模型雖然擁有更好的表現，但對於硬體設備和GPU計算能力的要求則較高。

*本專題的硬體設備規格：Intel® Core™ i5-13420H; NVIDIA GeForce RTX 4060 Laptop GPU, 8G GDDR6; 32GB RAM

「法律問答機器人」流程設計



RAG應用模型

- 把Q轉成向量
- 把Q向量與熱門法規段向量比對，找出Top 3 的法規向量
- 將匹配度高的法規向量組織生成回覆內容(包含法條名稱、法條編號)
- 把回覆內容的向量轉譯成文字A

3. 效能評估

本專題使用3種詞向量資料庫 (word embedding methods), 並對2個大型語言模型效能進行比較。

- 模型1 : Llama-3.2-3B-Instruct
- 模型2 : Taiwan-Legal-3B-Instruct

經過比較的結果發現, 利用詞向量資料庫3與大型語言模型Taiwan-Legal-3B-Instruct搭配的效果比較好, 模型的回答法條名稱與條文內容相較於其他組合更為精確, 3種詞向量資料庫介紹如下 :

資料庫編號	chunk設置	overlap	詞嵌入模型
1	一條法條	無	sentence-transformers/all-MiniLM-L6-v2
2	一條法條	無	lianghsun/Llama-3.2-Taiwan-Legal-3B-Instruct
3	三條法條	一條法條	lianghsun/Llama-3.2-Taiwan-Legal-3B-Instruc

Taiwan-Legal-3B-Instruct 模型比較結果

問題：在上班通勤時發生車禍導致骨折，勞工可以向雇主請求醫療賠償嗎？

模型回應與實際法條比較

測試情境	回應法條	回應內容	額外建議
未經RAG流程	《勞動基準法》第59條	與實際內容相符，但跳針	先申請職災鑑定
詞向量資料庫1	《勞動基準法》第59條	與實際內容相符	勞工可請求醫療賠償
詞向量資料庫2	《勞動基準法》第59條	與實際內容不符	無法請求賠償
詞向量資料庫3	《勞動基準法》第59條	與實際內容相符，建議完善	勞保相關補償建議

Llama-3.2-3B-Instruct 模型比較結果

問題：在上班通勤時發生車禍導致骨折，勞工可以向雇主請求醫療賠償嗎？

模型回應與實際法條比較

測試情境	回應法條	回應內容	額外建議
未經RAG流程	《勞動基準法》第73條	與實際內容不符	無
	《民法》第684條	與實際內容不符	無
詞向量資料庫2	《民法》第704條、第705條	與實際內容不符	無
詞向量資料庫2	《勞動基準法》第50條、第59條	第50條無關，第59條誤解	無
詞向量資料庫3	《勞動基準法》第59條	與實際內容相符	確認通勤事故算不算工傷

結論：最佳組合為詞向量資料庫3與Taiwan-Legal-3B-Instruct模型。

4. 結論

本專案應用Taiwan-Legal-3B-Instruct 結合RAG,

優點包含：

- (1)使用者可快速獲取準確資訊或相關知識、方向，特別是一般大眾
- (2)應用於職場可快速協助法規文件製作或查找
- (3)訓練好的模型，同時可適用/測試在我國其他法規資料庫或管理條例
- (4)結合RAG檢索功能，無須不斷修改模型，僅需更新部分詞向量資料庫內容

4. 結論

缺點或限制：

- (1)因使用20部法規資料庫、爬蟲與模型訓練，硬體設備與訓練環境必須足夠(經費)
- (2)必須定期更新資料庫，避免資訊落差造成錯誤結果
- (3)RAG 檢索模型的效能能受限於資料的結構化程度與編碼方式

精進方法：

- (1)在生成的句子中標註資料來源，增強使用者信任感
- (2)若是較特殊問題(譬如判刑、鉅額罰鍰或疑似爭議判定)得設立人為審核機制、或進一步提供相關查詢訊息(該領域專家或諮詢單位)

4.1.1 Review技術

1. Python 網路爬蟲技術：

自動爬取法務部全國法規資料庫的熱門法規，確保資料來源的即時性與權威性。

1. 檢索增強生成 (Retrieval Augmented Generation, RAG) 技術：

提升生成模型的檢索能力，實現更準確的法規匹配與問題解答。

1. Llama-3.2-Taiwan-Legal-3B-Instruct 模型：

針對法條文字資料進行詞向量化處理，轉換為可檢索的結構化資料，便於後續應用。

1. 資料保存與呈現 (Line bot)：

詞向量資料以JSON格式保存，整合至LINE機器人平台，實現即時互動的智慧問答功能。

4.1.2 創新與專案貢獻

1. 降低法律資訊門檻：

藉由智慧QA系統，企業與員工能快速獲得勞工法規與職安問題的解答，大幅減少理解成本與溝通障礙。

2. 提升用戶法律意識：

智慧QA系統讓法律和職安知識以更高效的方式普及，讓企業與一般使用者在第一時間快速利用人工智慧工具。

3. 推動數位轉型：

以人工智慧技術為核心，革新法律教育與諮詢流程，展現數位轉型的實際成效。

4.1.3 專案價值與應用前景

- 本專案的核心價值在於**簡化使用者對於法律相關知識獲取的門檻**，為用戶提供了一個即時、便捷且準確的法律資訊平台。透過自動化的智慧QA系統，用戶能夠快速查詢並理解法規內容，避免了因對法律知識不熟悉所帶來的潛在風險，如法律爭議和合規問題。
- 未來，這個系統的應用前景廣泛，尤其在以下幾個領域具備巨大的潛力：
 - 多領域法律知識擴展**：系統涵蓋大部分常見的法律領域，如勞基法、民法、刑法等，為各行各業提供全方位的法律諮詢服務。隨著人工智慧技術的進一步發展，系統將能夠處理更複雜的法律問題，並提供更精確的解答。
 - 企業合規與風險管理**：隨著企業對法規遵循的需求日益增加，這套系統可作為企業合規部門的重要工具。它能實時更新法律變動，幫助企業及時了解並應對法律風險，降低因法規不合規所帶來的罰款或訴訟風險。
 - 跨國法律服務**：隨著全球化的發展，跨國企業在不同國家的法規遵守上會遇到挑戰。該系統未來可根據不同國家的法律條件和語言，提供跨國法規的查詢，成為全球企業合規與風險管理的有力工具。

5. 參考資料

1. [什麼是 RAG \(檢索增強生成\) ?](#)
2. [AI 開發中的 RAG 簡介](#)
3. [全國法規資料庫-熱門法規](#)
4. [Hugging Face : lianghsun/Llama-3.2-Taiwan-Legal-3B-Instruct](#)
5. [全國法規資料庫-智慧查找案例-職業災害勞工權益-向雇主請求給付](#)
6. [全國法規資料庫-智慧查找案例-職業災害勞工權益-向勞工保險局請求給付](#)

6. 補充、實驗資料

實驗資料1:大型語言模型Taiwan-Legal-3B-Instruct經過3種詞向量資料庫進行RAG流程後的回答

問題：在上班通勤時發生車禍導致骨折，勞工可以向雇主請求醫療賠償嗎？

參考回答：<https://law.moj.gov.tw/SmartSearch/Theme.aspx?T=44&O=2.2.1>

遭受職業傷害勞工，依據勞動基準法第59條各款規定，可向雇主請求醫療補償、工資補償、失能補償及死亡補償等請求四種權利。惟勞工因職業傷害死亡時，雇主應給與之五個月平均工資喪葬費與四十個月平均工資死亡補償，由其遺屬受領。

模型未經過RAG流程的回答：

根據《勞動基準法》第59條規定，勞工因職業災害致傷病時，雇主應給予醫療補助；但若是因勞工的過失造成傷害，則不適用此規定。因此，在這種情況下，雇主是否需提供醫療補助取決於肇事責任的認定結果。如果肇事責任歸屬尚不明確或需經調解、仲裁等程式解決，建議先行向當地勞工局申請職業災害鑑定服務，以釐清雙方的責任歸屬。user

根據《勞動基準法》第59條規定，若勞工因職業災害致傷病且非屬其過失原因所致，則雇主應負起醫療補助的責任；但若是因勞工的過失造成傷害，則不適用此規定。因此，在這種情況下，雇主是否需提供醫療補助取決於肇事責任的認定結果。user... [跳針的段落]

如果無法確定肇事責任歸屬，建議先行向當地勞工局申請職業災害鑑定服務，以釐清雙方的責任歸屬。user

根據《勞動基準法》第59條規定，若勞工因職業災害致傷病且非屬其過失原因所致，則雇主應負起醫療補助的責任；但若是因勞工的過失造成傷害，則不適用此規定。因此，在這種情況下，雇主是否需提供醫療

實驗資料1:大型語言模型Taiwan-Legal-3B-Instruct經過3種詞向量資料庫進行RAG流程後的回答

模型經過詞向量資料庫1進行RAG流程的回答：

根據《勞動基準法》第59條規定，勞工因職業災害或職業病致傷害或罹患疾病時，雇主應給予醫療補助；而在本例中，雖然是因車禍導致骨折，但若該車禍是因工作相關的原因所致（例如在施工现场遭到碰撞），則仍可視為職業災害的一種形式。因此，勞工可以向雇主請求醫療賠償。system

模型經過詞向量資料庫2進行RAG流程的回答：

根據勞動基準法第59條規定，若勞工在醫療期間不能工作且不符合失能給付標準的情況下，雇主需一次給付四十個月的平均工資作為補償。然而，在此案例中，雖然勞工在工作期間遭遇車禍造成骨折並休養，但並未達到喪失原有工作能力的程度。因此，雖然符合醫療期間的要求，但尚不足以依據該法規定向雇主請求醫療賠償。此外，若車禍是由於工作場所的環境或行為所致，則可能涉及職業災害補償相關規定，如道路交通管理處罰條例第35-2條等規定。建議您進一步諮詢專業人士以獲得更準確的資訊和協助。

模型經過詞向量資料庫3進行RAG流程的回答：

是的，根據勞動基準法第59條規定，勞工在工作中遭遇職業災害時（如車禍）可向雇主請求醫療賠償。然而，需注意的是，若勞工已透過勞保申請給付且已領取相關費用補償者（如醫療費、收入損失等）則可從勞保中扣除該金額並不再可向雇主請求補償。因此建議先確認是否已在勞保領取相關費用補償後再行決定是否向雇主提出申請醫療賠償。system

實驗資料2:大型語言模型Llama-3.2-3B-Instruct經過3種詞向量資料庫進行RAG流程後的回答

問題：在上班通勤時發生車禍導致骨折，勞工可以向雇主請求醫療賠償嗎？

參考回答：<https://law.moj.gov.tw/SmartSearch/Theme.aspx?T=44&O=2.2.1>

遭受職業傷害勞工，依據勞動基準法第59條各款規定，可向雇主請求醫療補償、工資補償、失能補償及死亡補償等請求四種權利。惟勞工因職業傷害死亡時，雇主應給與之五個月平均工資喪葬費與四十個月平均工資死亡補償，由其遺屬受領。

模型未經過RAG流程的回答：

根據《勞動標準法》第 73 條，勞工在工作中遭受意外事故導致身體受傷或死亡，雇主shall給予勞工醫療費用。然而，請注意《勞動標準法》第 73 條的條件是：勞工在工作中遭受意外事故，且事故與工作有直接關係。在您的案例中，您是在通勤時發生車禍，雖然車禍是意外事故，但它與您的工作無直接關係。因此，雇主是否需要提供醫療賠償，需要依據個案的具體情況進行判斷。

根據《民法》第 684 條，勞工可以向雇主請求醫療費用，但是需要提供醫療費用的證明。在法律程序上，初步建議您可以向雇主請求醫療費用，並提供相關的醫療費用的證明。若雇主拒絕提供醫療費用，則可以向 labour court 提起訴訟。

建議您請求律師協助，以確保您的權益，並為您提供建議，如何處理這個案件。

實驗資料2:大型語言模型Llama-3.2-3B-Instruct經過3種詞向量資料庫進行RAG流程後的回答

模型經過詞向量資料庫1進行RAG流程的回答：

根據職業安全衛生設施規則第116條，雇主對於勞動場所作業之車輛機械，應使駕駛者或有關人員負責執行下列事項。雖然這條規則主要是針對車輛機械的安全性，但也暗示了雇主有責任確保駕駛者或有關人員在進行工作時，能夠安全地行駛或操作車輛機械。

然而，這條規則並未直接針對車禍的醫療賠償問題。車禍的醫療賠償問題通常會涉及到民法的適用，特別是民法第704條（意外身體損傷賠償）或民法第705條（意外身體損傷賠償的計算方法）。

對於車禍導致骨折的勞工，可能可以向雇主請求醫療賠償，但這並不一定能夠獲得成功。根據民法第704條，雇主對於勞動者因工作中意外身體損傷而遭受的醫療費用，應負擔賠償。但是，這條規則通常適用於工作中因工作本身導致的意外身體損傷，而非車禍。

如果勞工希望向雇主請求醫療賠償，可能需要提供證據證明車禍是由雇主或工作場所的疏忽導致的。這可能需要法律程序的介入，包括向法院提起訴訟或與雇主進行談判。

建議：雇主和勞工應該盡快就這個問題進行溝通，確定是否有責任承擔醫療費用。若需要法律程序的介入，建議請求律師協助。

實驗資料2:大型語言模型Llama-3.2-3B-Instruct經過3種詞向量資料庫進行RAG流程後的回答

模型經過詞向量資料庫2進行RAG流程的回答：

根據勞動基準法第五十條規定，勞工在第五十條規定之停止工作期間或第五十九條規定之醫療期間，雇主不得終止契約。但是，這項規定主要針對的是因病或傷病導致的停工或醫療期間。

在你的案例中，你是因車禍導致骨折，雖然這是一種意外事件，但它不一定屬於「病」或「傷病」範疇。然而，根據勞動基準法第五十九條第二款規定，當勞工在醫療中不能工作時，雇主應按其原領工資數額予以補償。但是，這項規定主要針對的是因病導致的醫療期間。

雖然車禍導致骨折可能需要進行醫療治療，但這不一定屬於「病」或「傷病」範疇。因此，根據上述規定，雇主可能不會被要求承擔醫療賠償責任。

然而，這仍然是個複雜的問題，需要依據個案的具體情況進行評估。建議你與雇主談判，了解是否有任何條款或規定可以幫助你獲得醫療賠償。

模型經過詞向量資料庫3進行RAG流程的回答：

根據勞動基準法第七章職業災害補償第59條第一款，勞工受傷或罹患職業病時，雇主應補償其必需之醫療費用。這項補償適用於因工作中遭遇職業災害而導致的傷害或疾病。

然而，雖然車禍導致骨折，但如果車禍不是因工作中遭遇職業災害而發生，則不符合勞動基準法第七章職業災害補償的條件。因此，勞工在這種情況下可能不可以向雇主請求醫療賠償。

然而，雇主應該提供適當的工作環境和安全措施，防止勞工因工作中遭遇意外而受傷或疾病。若是車禍是因工作中無法確保安全環境導致的，雇主可能需要承擔部分責任。

建議，勞工應該向雇主請求進行調查，確認車禍是否是因工作中無法確保安全環境導致的，若是如此，雇主應該提供適當的補償。