



Slumber Alert



— 基於**HRV**進行睡眠警示

第六組

葉霽恩 / 林軒如 / 黃宇成 / 陳芃瑋 / 崔馨予 / 魏榛均 / 張敏奕



Outline.

01 題目選擇緣由

02 產品概念介紹

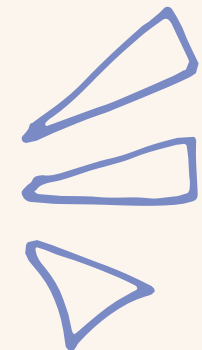
03 功能及技術盤點

04 研究資料

05 組員分工

06 時程總結回顧

07 成果



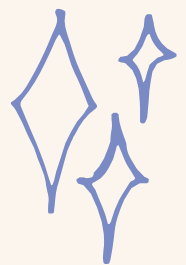


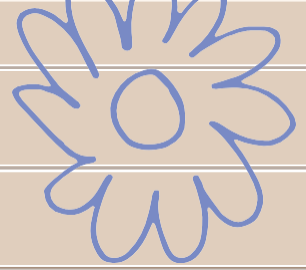
01 題目選擇緣由

□ 開發低成本/高精準HRV睡眠狀態預測模型

□ 達成SDGs 目標3

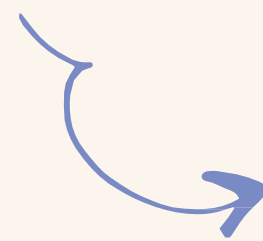
→ 確保及促進各年齡層健康生活與福祉

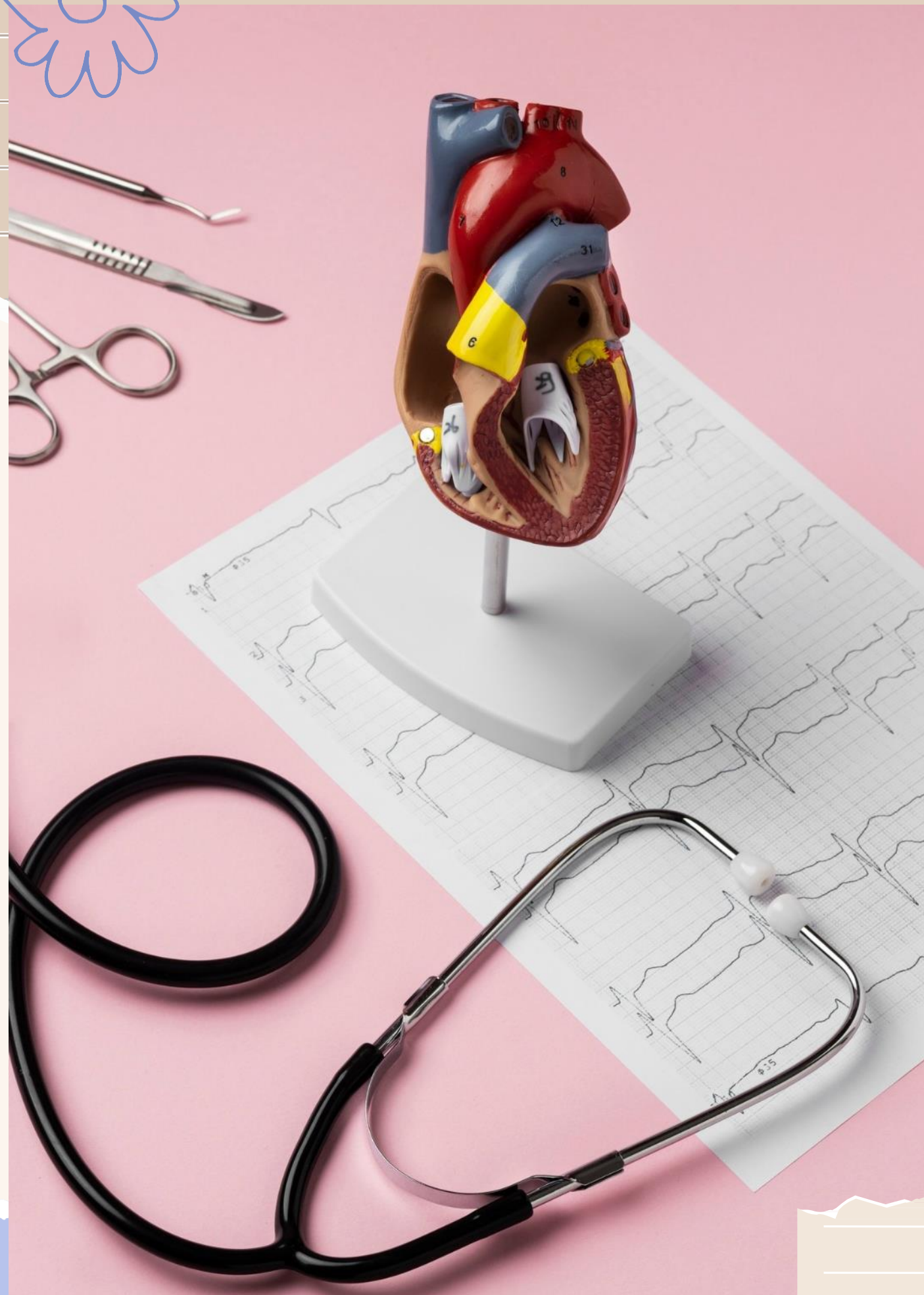




02 產品概念介紹

- 使用HRV預測睡眠狀態，並以此了解自身睡眠表現





03 功能及技術盤點

□ 資料預處理

□ 隨機森林

□ SVM

□ 決策樹

□ 使用者介面

資料預處理

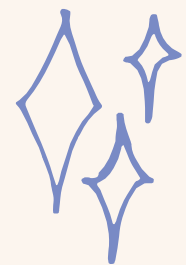
You Snooze, You Win: the PhysioNet/Computing in Cardiology Challenge 2018

Clinical Feature	Total	Training	Test
Sample size	1,983	994	989
Age	55 (14.4)	55 (14.3)	55 (14.4)
Gender (% male)	65	67	63

Time spent in sleep stage (%)

Wake	29.3	28.0	31.0
NREM 1	19.5	19.6	19.0
NREM 2	51.3	51.0	51.7
NREM 3	13.8	14.0	13.8
REM	15.3	15.5	15.2

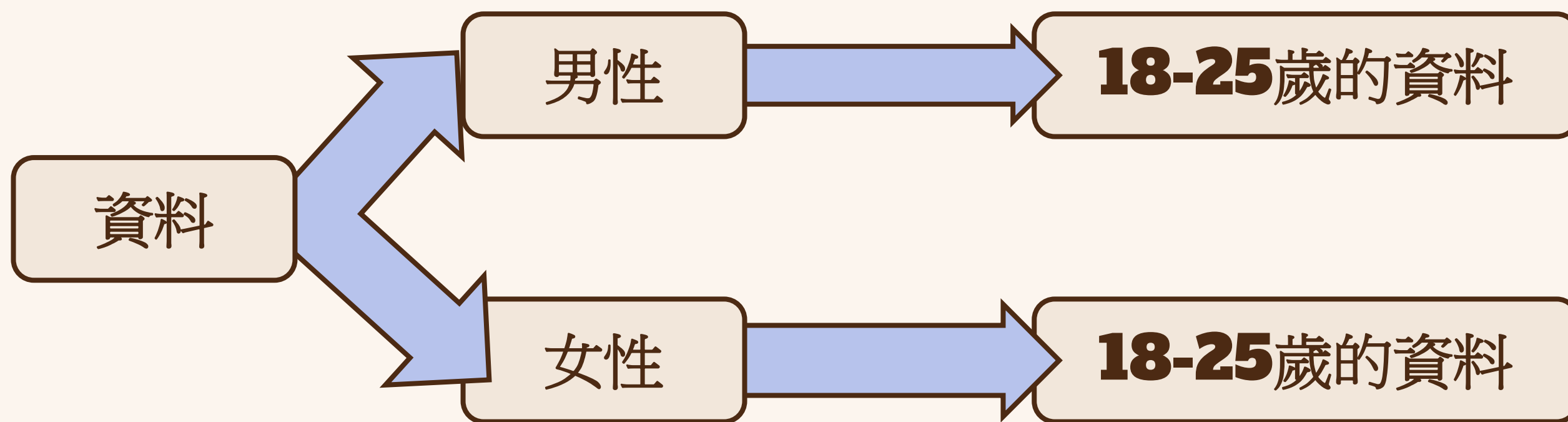
→我們有的資料及來源



資料分類

→我們將資料進行分類

→希望能將個體差異極小化



18-25female_data

18-25male_data

tr03-0645

tr03-1155

tr03-1333

tr04-0468

tr04-0829

tr05-0356

tr05-0864

tr06-0332

tr08-0137

#bbi計算函式

def bbi_cal(sleep_stage, i):

ecg_to_cal = sleep_stage[i]

ecg_biosppy_out = biosppy.signals.ecg.ecg(signal=ecg_to_cal, show=False, sampling_rate=record.__dict__.get('fs'))

ts_rpeaks = ecg_biosppy_out.__dict__.get('ts')[ecg_biosppy_out.__dict__.get('rpeaks')]

bbi = (ts_rpeaks-np.roll(ts_rpeaks, 1))[1:]*1000

bbi_list = bbi.tolist()

for j in bbi_list:

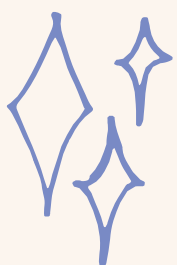
 if j>2500:

 bbi_list.remove(j)

bbi=np.array(bbi_list)

return bbi

→將讀取進來的資料進行異常值處理，並進行bbi計算




```
def hrv_cal(bbi):
    sdn = []
    rmssd = []
    pnn50 = []
    hf = []

    for k in range(0, len(bbi), 60):
        if k + 60 <= len(bbi):
            group_bbi = bbi[k:k+60]
        else:
            group_bbi = bbi[k:]

        # 計算sdnn
        sdn.append(td.sdn(nni=group_bbi)['sdnn'])
        # 計算rmssd
        rmssd.append(td.rmssd(nni=group_bbi)['rmssd'])
        # 計算pnn50
        nn50_res = td.nn50(nni=group_bbi)
        pnn50.append(nn50_res['pnn50'])
        # 計算hf
        freq_results = fd.welch_psd(nni=group_bbi)
        hf.append(freq_results['fft_norm'][1]) # 取HF值

    sdn, rmssd, pnn50, hf = np.array(sdn), np.array(rmssd), np.array(pnn50), np.array(hf)
    #print(pnn50)

    print(len(sdn), len(rmssd), len(pnn50), len(hf))

    return sdn, rmssd, pnn50, hf
```

→將bbi計算數值進行hrv計算
(sdnn,rmssd,nn50,hf)



SDNN（**NN**間期的標準差）：

較高的SDNN值通常被認為代表自主神經系統整體活動較高，功能狀態較好。健康人群的SDNN值較高，而患有某些疾病(如糖尿病、心力衰竭等)的人群的數值常常降低。

RMSSD（相鄰**NN**間期的均方根差）：

相鄰心搏間期的均方根差，通常以毫秒為單位。較大的RMSSD值表示心臟跳動的快速變化，反映了副交感神經活動的增加。



pNN50 :

相鄰NN間期差異大於50毫秒的百分比。

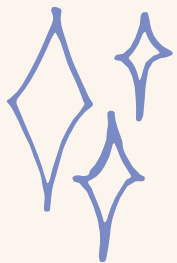
較高的pNN50值表示心臟跳動的不規則性較大，自主神經系統功能較好。

HF（高頻）：

頻域指標，用於分析不同頻率範圍內的心率變異。LF反映交感神經活動，而HF反映副交感神經活動。它們的比值（LF/HF）可以提供有關自主神經系統平衡的信息。

將所有的參數寫入csv檔案方便訓練模型
年齡、性別、睡眠狀態、
bbi,sdnn,rmssd,pnn50,hf

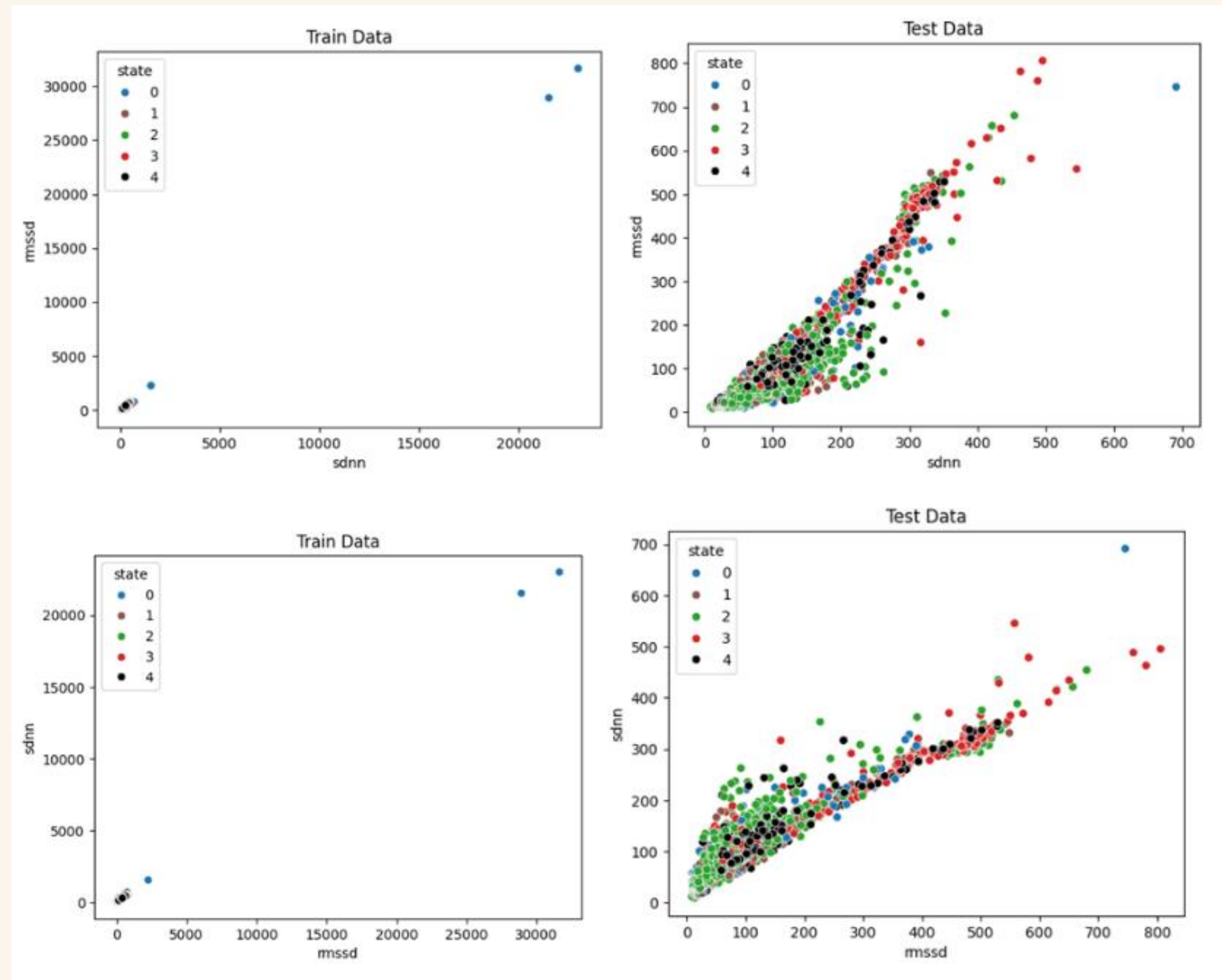
```
#####  
#csv資料架構  
# data[][] = [  
#0 1 2      [record_name, gender , age,  
#3 4        'status', 睡眠狀態紀錄,  
  
#5 6        'wake_bbi',wake_bbi,  
#7 8        'rem_bbi',rem_bbi,  
#9 10       'n1_bbi',n1_bbi,  
#11 12      'n2_bbi',n2_bbi,  
#13 14      'n3_bbi',n3_bbi,  
  
#15 16      'wake_sdnn',wake_sdnn,  
#17 18      'rem_sdnn',wake_sdnn,  
#19 20      'n1_sdnn',n1_sdnn,  
#21 22      'n2_sdnn',n2_sdnn,  
#23 24      'n3_sdnn',n3_sdnn,  
  
#25 26      'wake_nmssd',wake_nmssd,  
#27 28      'rem_nmssd',wake_nmssd,  
#29 30      'n1_nmssd',n1_nmssd,  
#31 32      'n2_nmssd',n2_nmssd,  
#33 34      'n3_nmssd',n3_nmssd,  
  
#35 36      'wake_pnn50',wake_pnn50,  
#37 38      'rem_pnn50',rem_pnn50,  
#39 40      'n1_pnn50',n1_pnn50,  
#41 42      'n2_pnn50',n2_pnn50,  
#43 44      'n3_pnn50',n3_pnn50,  
  
#45 46      'wake_pnn50',wake_hf,  
#47 48      'rem_pnn50',rem_hf,  
#49 50      'n1_pnn50',n1_hf,  
#51 52      'n2_pnn50',n2_hf,  
#53 54      'n3_pnn50',n3_hf  
#      ]  
#####
```



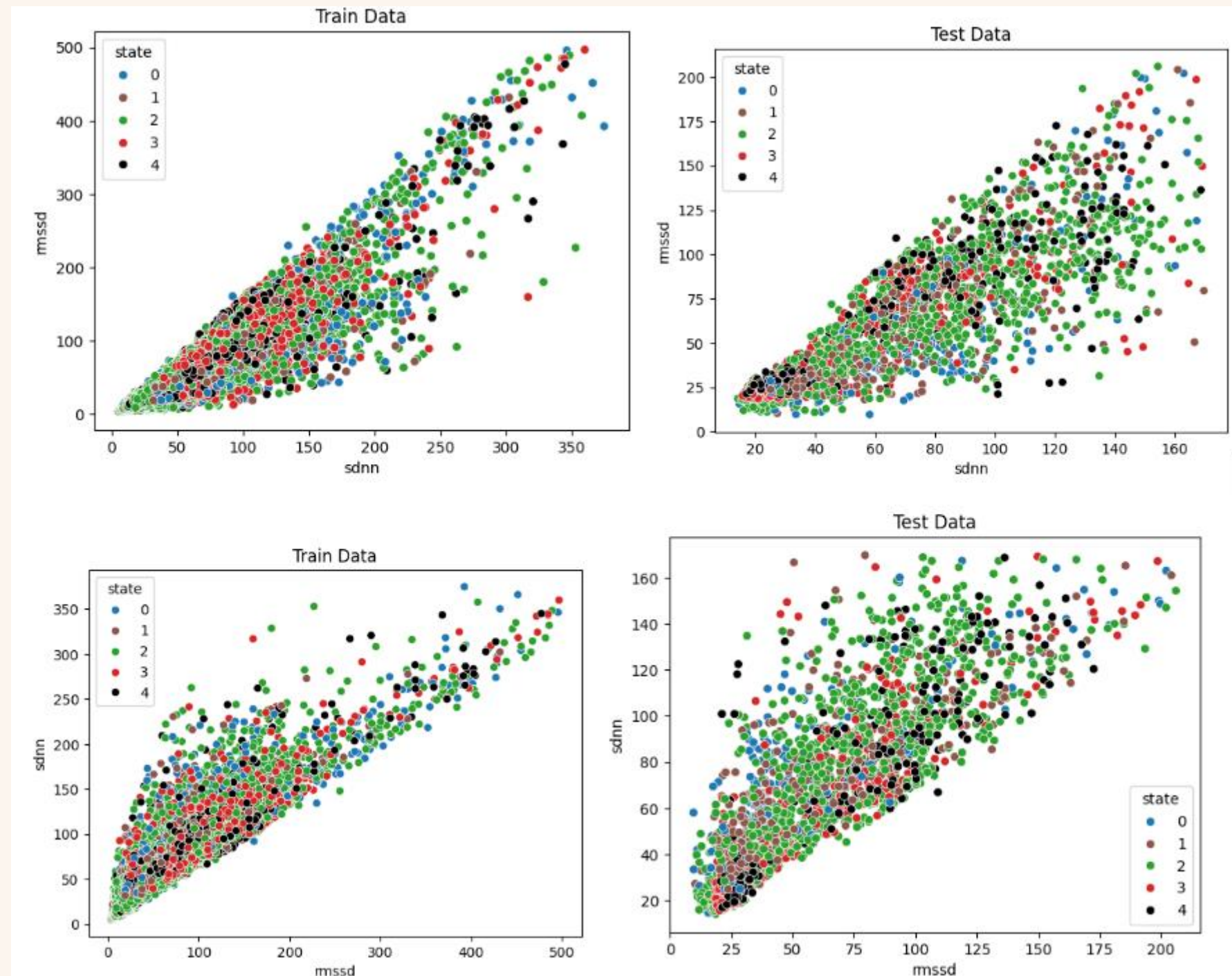
```
train_data = 11093
X_train = 10933
test_data = 2862
X_test = 2463
```

[illegible]

未過濾的資料sdnn-rmssd



過濾後的資料sdnn-rmssd

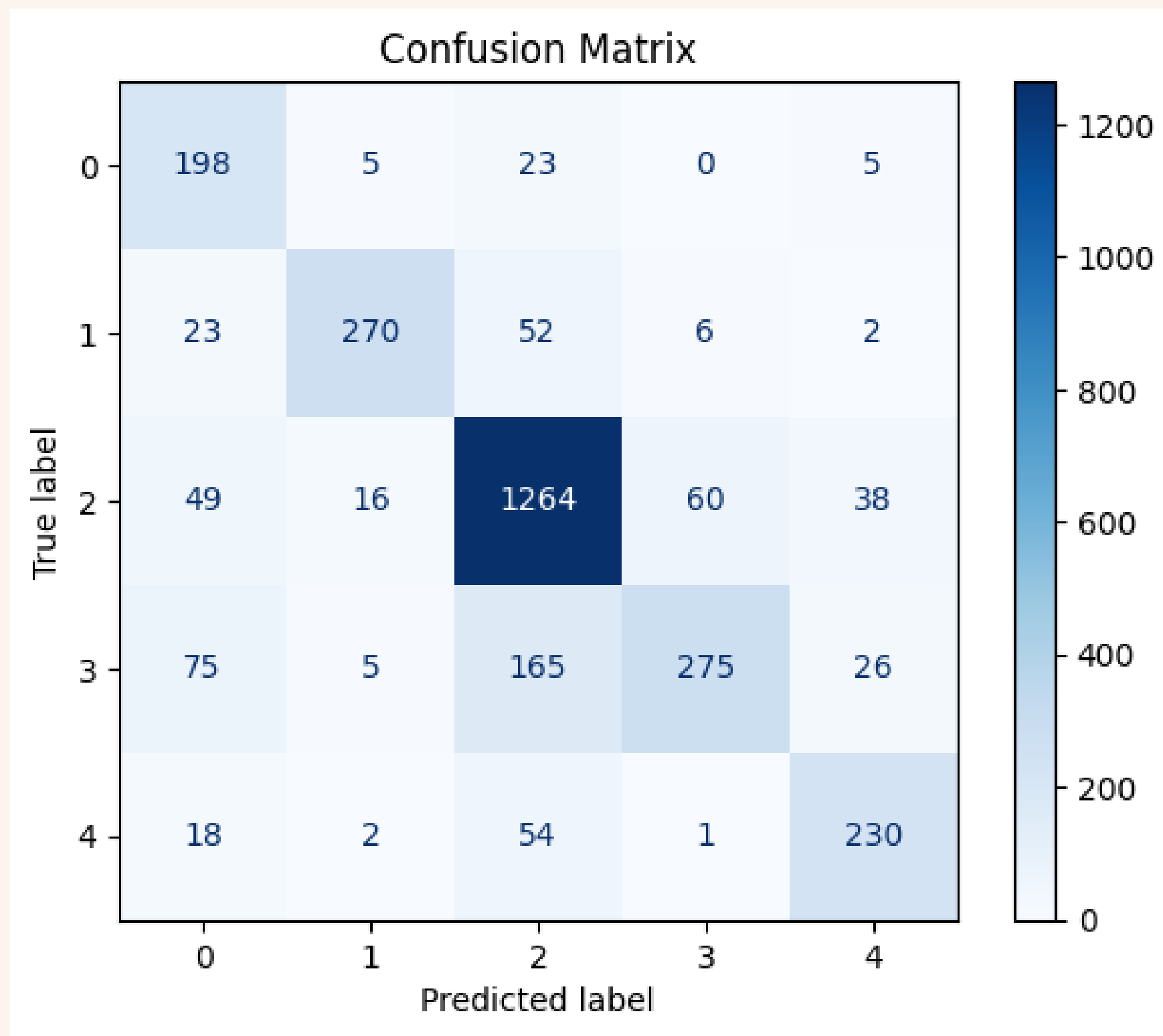


隨機森林

訓練結果(未過濾)

模型的準確度為： 0.7816212438853948

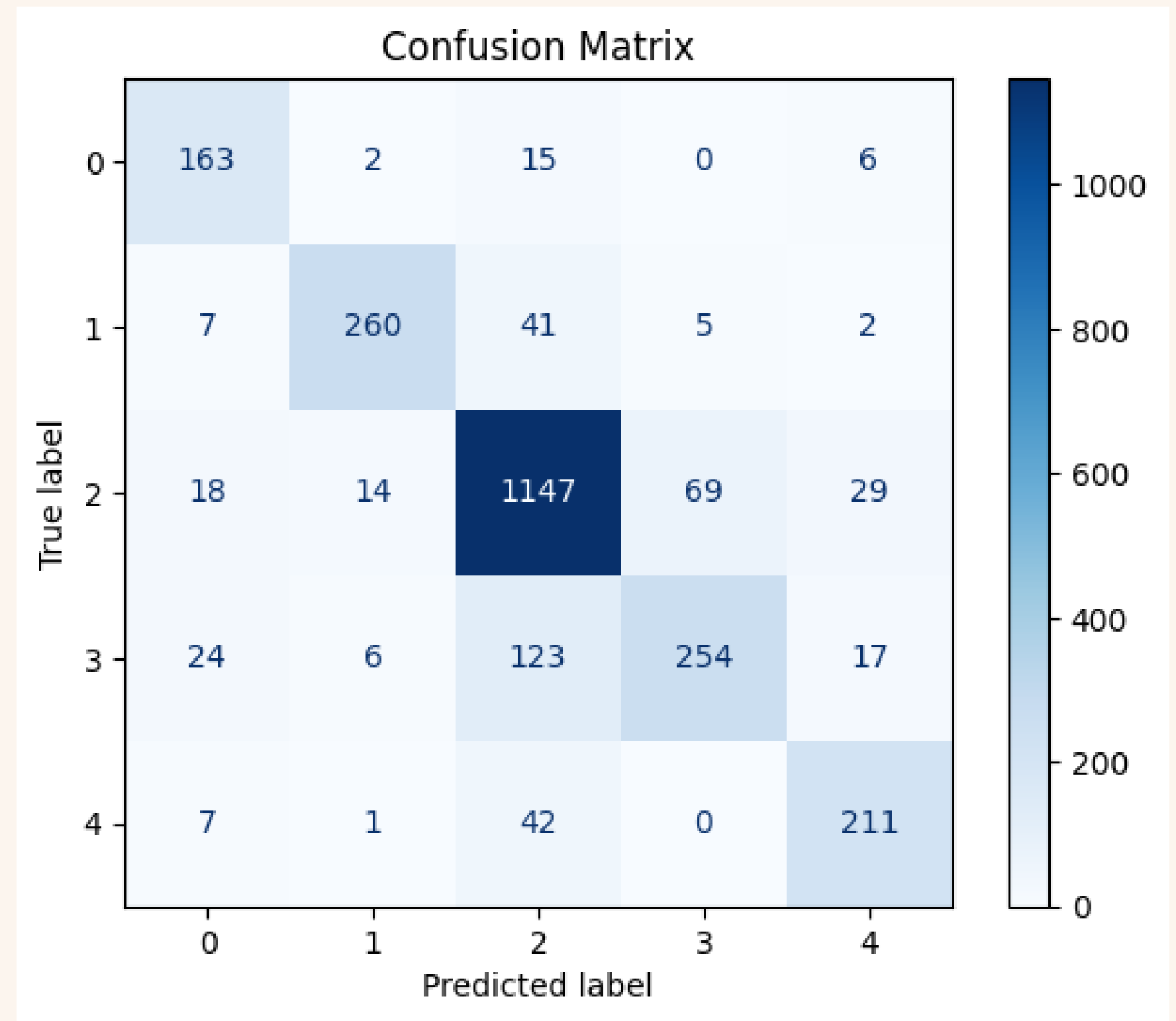
```
Prediction counts:  
State 0: 363 times  
State 2: 1558 times  
State 4: 301 times  
State 1: 298 times  
State 3: 342 times
```



訓練結果(過濾)

模型的準確度為： 0.8262281770198945

```
Prediction counts:  
State 0: 219 times  
State 4: 265 times  
State 1: 283 times  
State 2: 1368 times  
State 3: 328 times
```



SVM

→將資料標準化後

轉為pd.DataFrame格式

Preprocessing

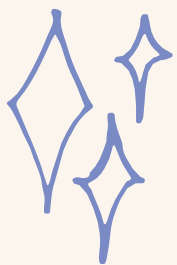
```
scaler = StandardScaler()  
X_scaled = scaler.fit_transform(X)  
X_scaled = pd.core.frame.DataFrame(X_scaled)  
X_scaled
```

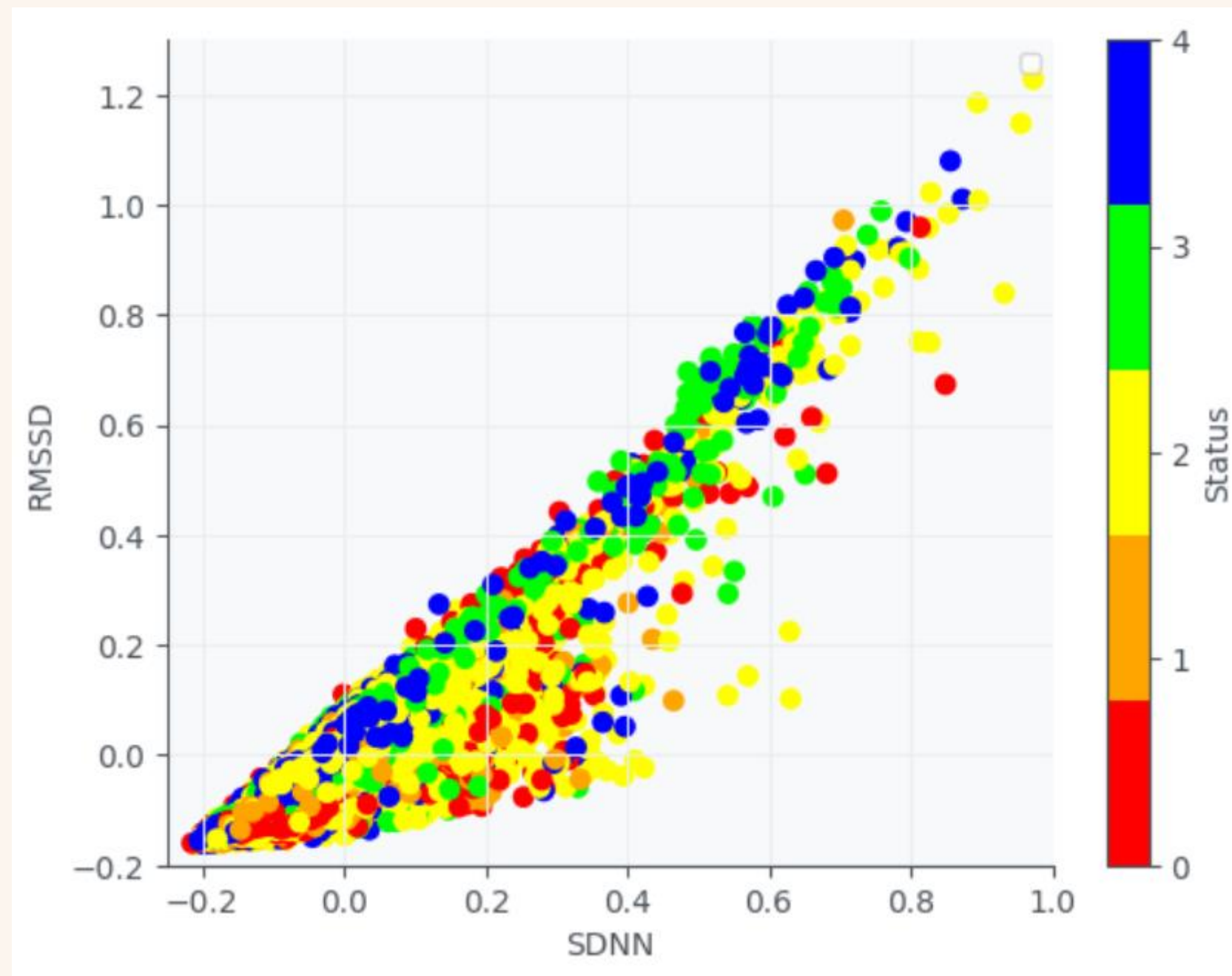
[55]

Python

```
...      0      1  
0  -0.051795 -0.105256  
1   0.323686  0.336965  
2   0.138980  0.176493  
3  -0.156231 -0.108625  
4  -0.053450 -0.095010  
...      ...      ...  
11251  0.184858  0.225799  
11252  0.578088  0.673608  
11253  0.585341  0.609714  
11254  0.598277  0.765607  
11255  0.856080  1.079740
```

11256 rows × 2 columns





RMSSD 與 SDNN以及標註睡眠狀態的散佈圖

資料重疊性高

(0=wake, 1=rem, 2=n1, 3=n2, 4=n3)

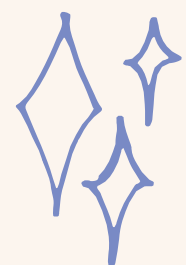
```
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test,y_pred))
print('\n')
print(classification_report(y_test,y_pred))
```

[93] Python

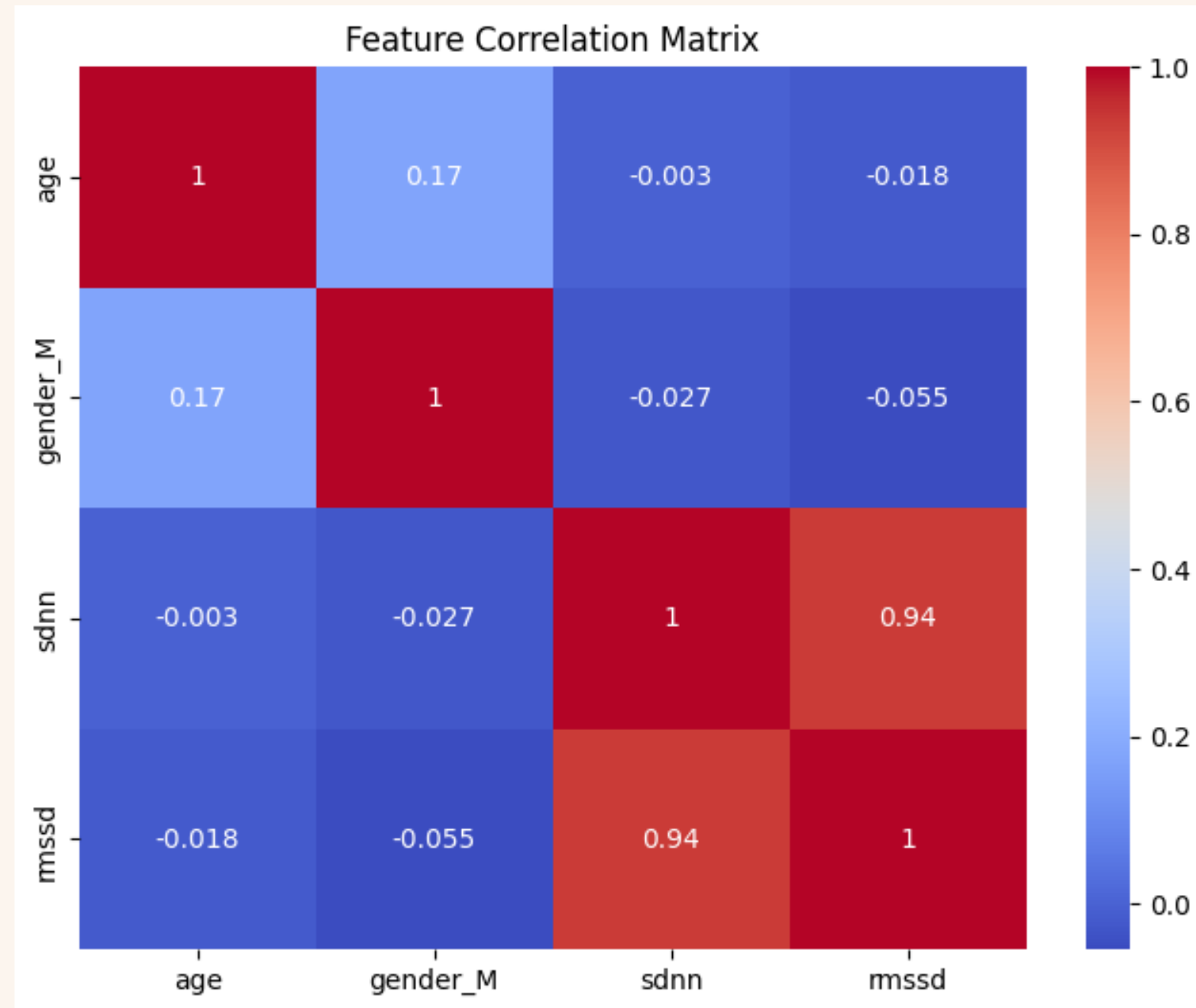
```
... [[ 0  0 228  2  0]
      [ 0  0 229  1  0]
      [ 0  0 1144 17  0]
      [ 0  0 377  8  0]
      [ 0  0 240  6  0]]
```

	precision	recall	f1-score	support
0.0	0.00	0.00	0.00	230
1.0	0.00	0.00	0.00	230
2.0	0.52	0.99	0.68	1161
3.0	0.24	0.02	0.04	385
4.0	0.00	0.00	0.00	246
accuracy			0.51	2252
macro avg	0.15	0.20	0.14	2252
weighted avg	0.31	0.51	0.36	2252

→ 準確率:0.51，如上的confusion matrix顯示模型預測偏向預測標記2(n1)、標記3(n2)



資料探索



age、gender、sdn、mssd :
→ 顯示數值特徵間相關性

決策樹

使用 **Z-Score** 做資料過濾

模型的準確度為： 0.2614660917287338

	precision	recall	f1-score	support
0	0.17	0.15	0.16	312
1	0.12	0.16	0.14	297
2	0.52	0.36	0.42	1324
3	0.08	0.12	0.10	344
4	0.13	0.20	0.15	274
accuracy			0.26	2551
macro avg	0.20	0.20	0.19	2551
weighted avg	0.33	0.26	0.29	2551

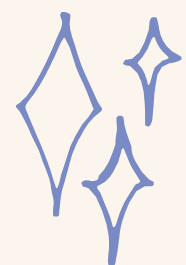
模型的準確度為： 0.2676332288401254

	precision	recall	f1-score	support
0	0.17	0.17	0.17	313
1	0.12	0.16	0.14	297
2	0.53	0.37	0.44	1324
3	0.09	0.13	0.10	344
4	0.12	0.17	0.14	274
accuracy			0.27	2552
macro avg	0.20	0.20	0.20	2552
weighted avg	0.33	0.27	0.29	2552

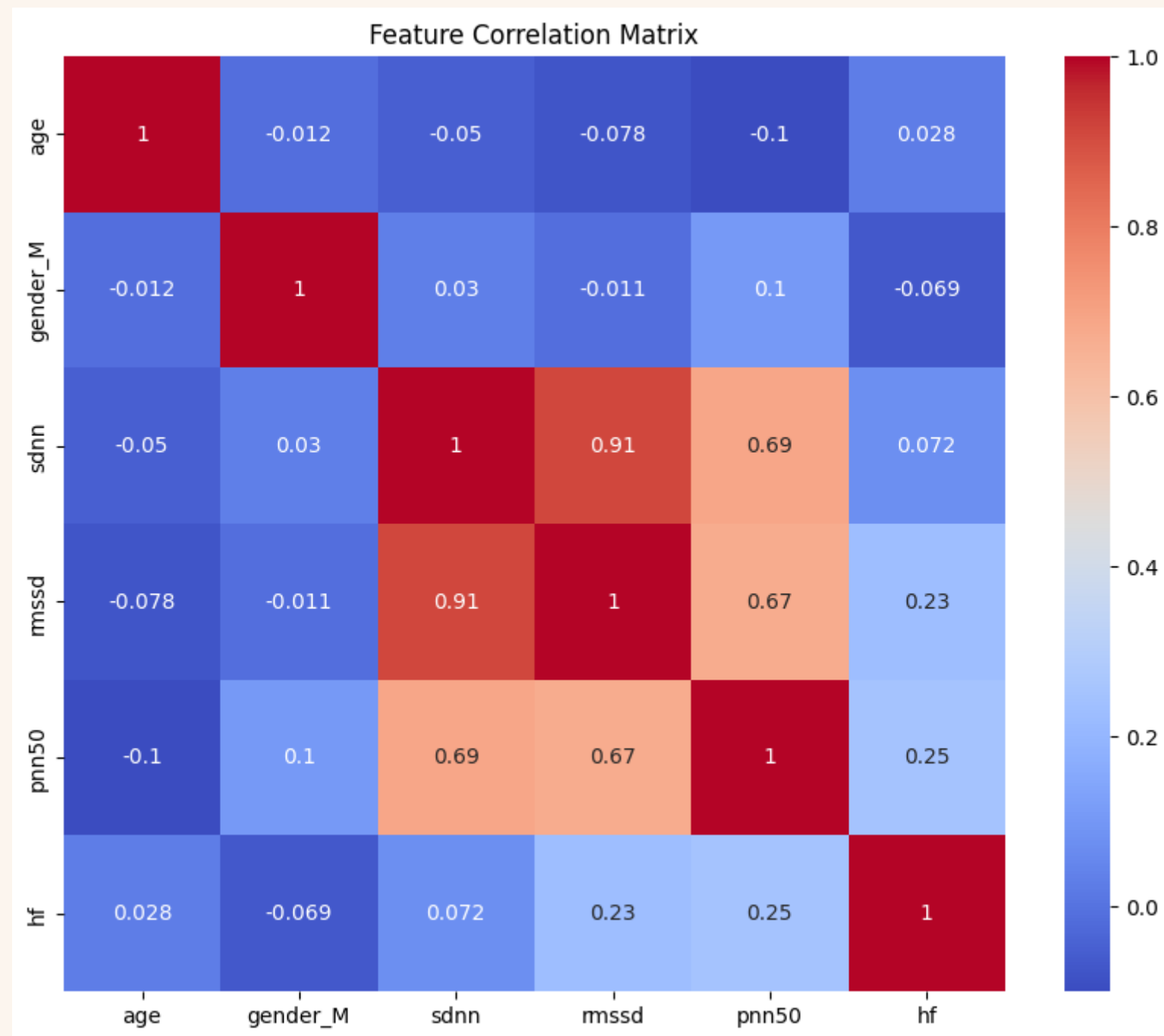
訓練結果(未過濾)

訓練結果(過濾後)

準確度無明顯提升



資料探索



新增 pnn50、hf 資料:
→顯示數值特徵間相關性

決策樹

使用 **Z-Score** 做資料過濾

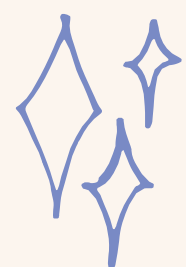
模型的準確度為： 0.7368421052631579				
	precision	recall	f1-score	support
0	0.81	0.52	0.63	417
1	0.76	0.84	0.80	434
2	0.87	0.77	0.82	1650
3	0.48	0.78	0.60	334
4	0.52	0.67	0.59	300
accuracy			0.74	3135
macro avg	0.69	0.72	0.69	3135
weighted avg	0.77	0.74	0.74	3135

模型的準確度為： 0.7649769585253456				
	precision	recall	f1-score	support
0	0.73	0.50	0.60	331
1	0.72	0.84	0.77	371
2	0.89	0.80	0.85	1394
3	0.47	0.82	0.60	272
4	0.80	0.72	0.76	236
accuracy			0.76	2604
macro avg	0.72	0.74	0.71	2604
weighted avg	0.80	0.76	0.77	2604

訓練結果(未過濾)

訓練結果(過濾後)

準確度提升約**3%**







使用者介面－使用**Django**開發

django



使用者介面－資料上傳



SDNN pNN50 Dropdown Search Something Search Hello, User! Dropdown

Your Age

22

Your Gender

Male

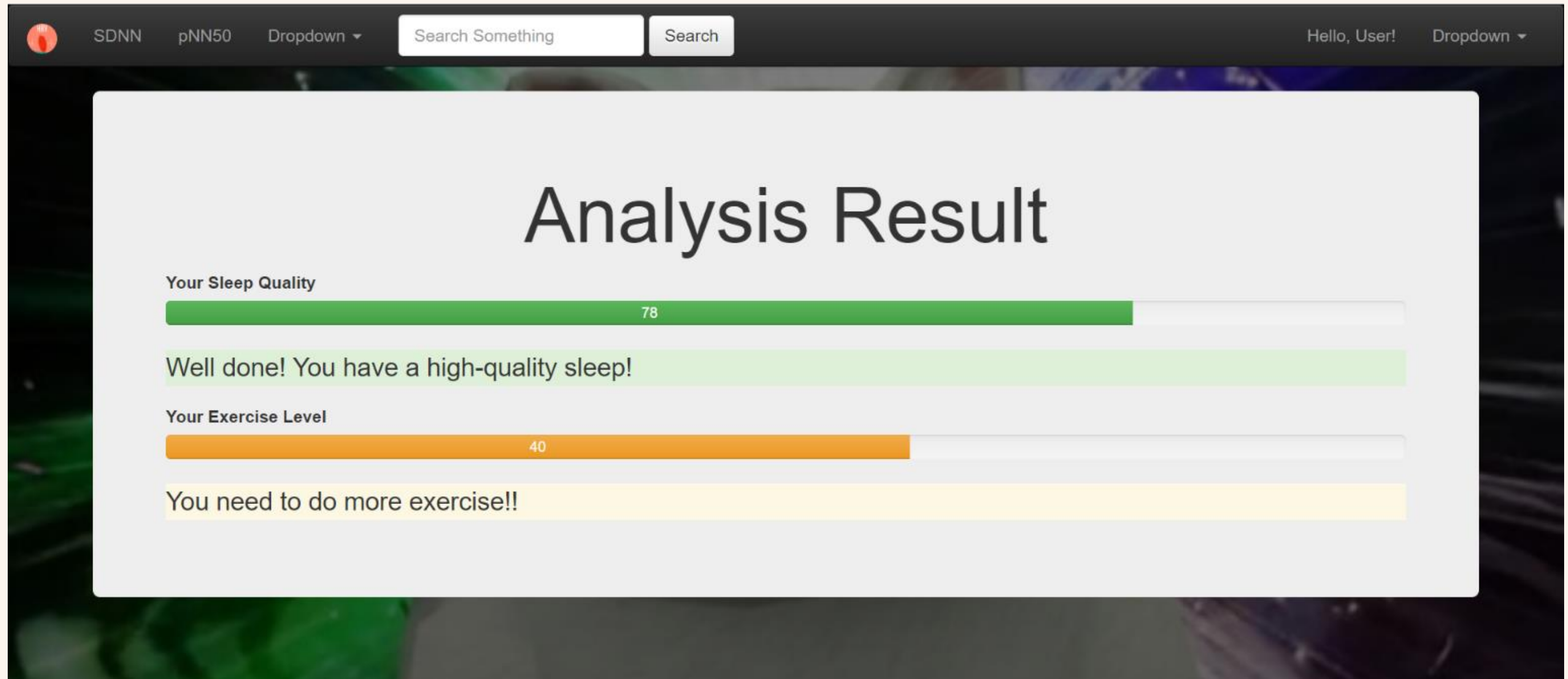
Your HRV File

選擇檔案 hrv_data.csv

Example block-level help text here.

Submit

使用者介面－結果分析



04 研究資料

文獻一

24小時時間域心率變異性與心率：
在**90**年間與年齡和性別關係

低心率變異性（**HRV**）被認為是許多病理狀況的標誌，包括死亡風險增加。
然而心率變異用於臨床預測時須考慮一些重要生理因素的影響。
其中最突出的有兩個因素：年齡和性別。

文獻二

健康受試者的心率變異性
與年齡和性別有關

通過分析 **101** 名健康受試者(**49** 名男性和 **52** 名女性)的 **24** 小時心電圖記錄，
研究年齡和性別對 **HRV** 的影響。
女性的心率變異性比男性稍低，即使在
年齡差異下也是如此。

05 組員分工

資料處理 (資料科學家)	黃宇成 崔馨予	介面呈現 (介面工程師)	葉霽恩
訓練模型 (模型工程師)	陳芃瑋 張敏奕 魏榛均	產品經理 (PM)	葉霽恩 林軒如



06 時程總結回顧

04/30

研究變因
建立模型

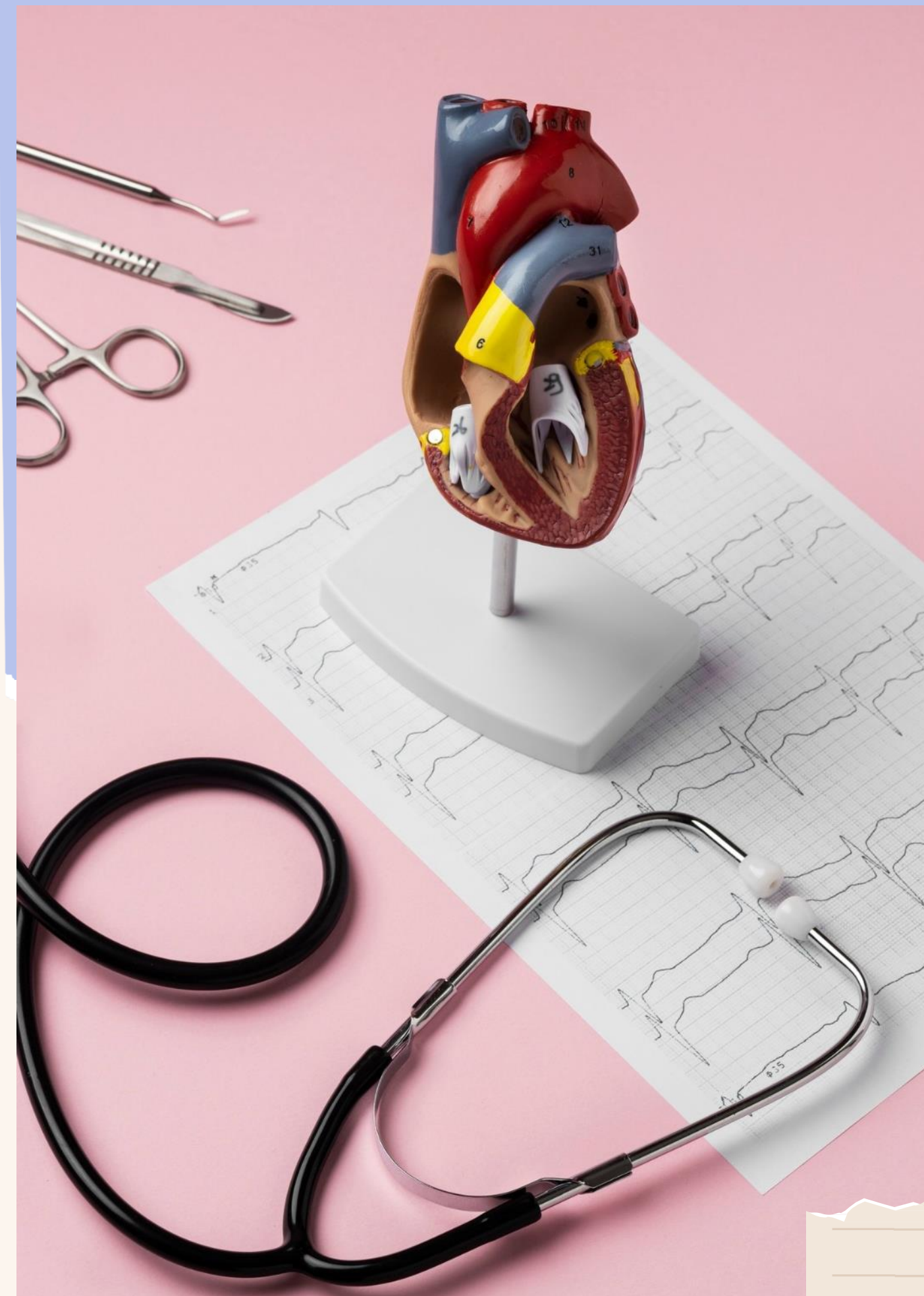
05/31

完成模型
測試及除錯



07 成果

- 利用隨機森林達成**82%**準確率的睡眠分期預測
- 使用**Django**開發使用者介面



Thanks!

