元智大學 工業工程與管理學系 畢業專題

使用統計分析篩選有效語音資料與特徵值以發展HIV感染者之情緒識別模型

YZUIEM 元智大學工業工程與管理學系

指導老師：林瑞豐 教授

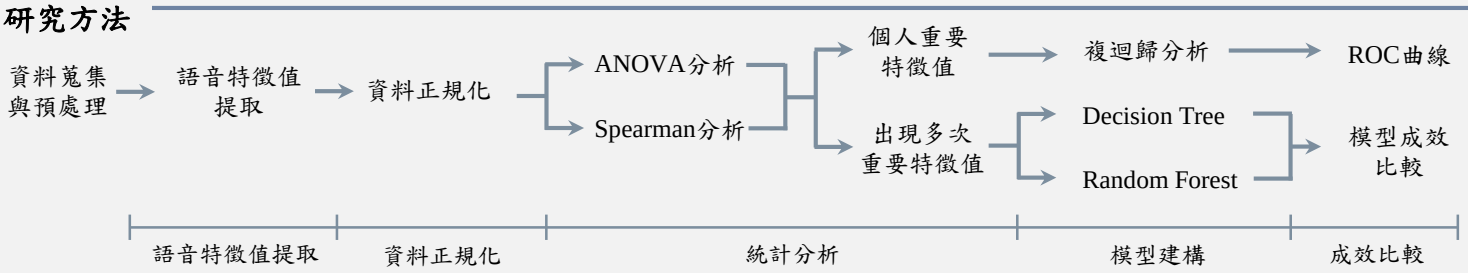
學生：潘巧育、陳亭妏

研究動機與目的

人們的情緒起伏會影響說話時的聲音表現，然而HIV感染者比一般人更常出現憂鬱或情緒低落之情況，但護理師及個案管理師無法及時照顧，因此我們想要藉由情緒會影響聲音的關係，來判斷患者的情緒是否低落。在吳允彤(2022)的研究中，使用元智大學工業工程與管理學系人因工程與設計實驗室與衛生福利部桃園醫院感染科合作期間蒐集之資料，進行HIV患者與物質成癮的語音情感辨識，發現資料有部分受測者回報不實之心情值與受測者回報之情緒分布不均兩個限制，導致在篩選重要特徵值時效果不佳，進而影響最後建模成效。

為了去除以上之資料限制，本研究將先著重於個別HIV受測者之聲音分析，透過針對個別受測者之資料進行單因子變異數分析(ANOVA)與相關性分析(Correlation)兩種不同統計分析，篩選顯著特徵值用於多種模型建構，預期能解決在吳允彤(2022)中所提之資料限制，獲得較好的情感辨識成效。

研究方法

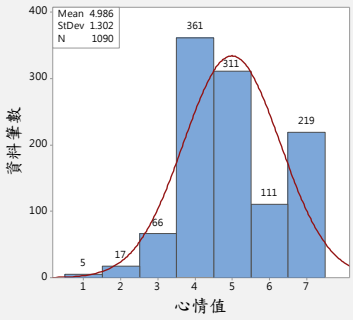


1) 39位受測者之錄音音檔
2) 利用Audacity去除雜音
3) 利用openSMILE提取聲音特徵值
4) 使用Jupyter Notebook將資料正規化
5) 使用ANOVA、Spearman分析篩選顯著特徵值
6) 使用決策樹和隨機森林建模
7) 模型成效比較
8) 使用複迴歸分析預測
9) ROC曲線

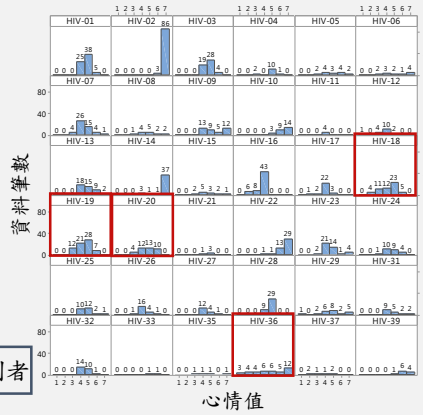
研究結果

受測者情緒分布

預處理後的總資料筆數



各受測者心情值分布圖



選用編號18、19、20、36之均勻受測者

出現多次之重要特徵值

顯著特徵值	累計次數	顯著特徵值	累計次數
pcm_fftMag_mfcc_sma_de[1]_linregerrQ	10	pcm_fftMag_mfcc_sma_de[3]_minPos	6
pcm_fftMag_mfcc_sma_de[1]_stddev	10	pcm_fftMag_mfcc_sma_de[3]_linregerrQ	6
F0_sma_de_minPos	9	pcm_fftMag_mfcc_sma_de[4]_max	6
pcm_fftMag_mfcc_sma_de[1]_minPos	9	pcm_fftMag_mfcc_sma_de[4]_minPos	6
pcm_fftMag_mfcc_sma_de[3]_stddev	9	pcm_fftMag_mfcc_sma_de[4]_linregc1	6
pcm_fftMag_mfcc_sma_de[5]_skewness	9	pcm_fftMag_mfcc_sma_de[4]_stddev	6
pcm_fftMag_mfcc_sma_de[9]_maxPos	9	pcm_fftMag_mfcc_sma_de[5]_min	6
jitterLocal_sma_de_minPos	9	pcm_fftMag_mfcc_sma_de[5]_linregerrQ	6
pcm_fftMag_mfcc_sma_de[1]_range	9	pcm_fftMag_mfcc_sma_de[9]_linregerrQ	6
pcm_fftMag_mfcc_sma_de[1]_linregc2	9	pcm_fftMag_mfcc_sma_de[10]_range	6
pcm_fftMag_mfcc_sma_de[2]_stddev	9	pcm_fftMag_mfcc_sma_de[12]_max	6
pcm_fftMag_mfcc_sma_de[4]_linregc2	9	pcm_fftMag_mfcc_sma_de[12]_amean	6
pcm_fftMag_mfcc_sma_de[4]_kurtosis	9	voiceProb_sma_de_minPos	6
pcm_fftMag_mfcc_sma_de[5]_stddev	9	F0_sma_de_amean	6
jitterLocal_sma_de_maxPos	9	jitterLocal_sma_de_linregc1	6
pcm_fftMag_mfcc_sma_de[2]_linregc1	9	shimmerLocal_sma_de_skewness	6
pcm_fftMag_mfcc_sma_de[2]_linregc2	9		

顯著特徵值	累積次數	顯著特徵值	累積次數
pcm_fftMag_mfcc_sma_de[2]_min	11	pcm_fftMag_mfcc_sma_de[12]_maxPos	7
pcm_fftMag_mfcc_sma_de[4]_linregc2	11	pcm_fftMag_mfcc_sma_de[12]_amean	7
voiceProb_sma_de_max	10	pcm_zcr_sma_de_kurtosis	7
pcm_RMSenergy_sma_de_minPos	9	pcm_RMSenergy_sma_de_maxPos	5
pcm_fftMag_mfcc_sma_de[1]_range	9	pcm_RMSenergy_sma_de_linregc1	5
pcm_fftMag_mfcc_sma_de[1]_minPos	9	pcm_RMSenergy_sma_de_kurtosis	5
pcm_fftMag_mfcc_sma_de[2]_linregerrQ	9	pcm_fftMag_mfcc_sma_de[1]_maxPos	5
pcm_fftMag_mfcc_sma_de[2]_stddev	9	pcm_fftMag_mfcc_sma_de[1]_linregc2	5
pcm_fftMag_mfcc_sma_de[3]_maxPos	9	pcm_fftMag_mfcc_sma_de[1]_kurtosis	5
pcm_fftMag_mfcc_sma_de[4]_linregc1	9	pcm_fftMag_mfcc_sma_de[3]_linregerrQ	5
pcm_fftMag_mfcc_sma_de[5]_min	9	pcm_fftMag_mfcc_sma_de[4]_skewness	5
pcm_fftMag_mfcc_sma_de[6]_amean	9	pcm_fftMag_mfcc_sma_de[4]_kurtosis	5
pcm_fftMag_mfcc_sma_de[6]_skewness	9	pcm_fftMag_mfcc_sma_de[5]_range	5
pcm_fftMag_mfcc_sma_de[11]_linregerrQ	9	pcm_fftMag_mfcc_sma_de[5]_skewness	5
voiceProb_sma_de_range	9	pcm_fftMag_mfcc_sma_de[8]_max	5
pcm_fftMag_mfcc_sma_de[1]_min	9	pcm_fftMag_mfcc_sma_de[8]_min	5
pcm_fftMag_mfcc_sma_de[1]_linregc1	9	pcm_fftMag_mfcc_sma_de[12]_linregerrQ	5
pcm_fftMag_mfcc_sma_de[2]_range	9	pcm_fftMag_mfcc_sma_de[12]_stddev	5
pcm_fftMag_mfcc_sma_de[3]_amean	9	pcm_zcr_sma_de_minPos	5
pcm_fftMag_mfcc_sma_de[4]_max	9	pcm_zcr_sma_de_linregc1	5
pcm_fftMag_mfcc_sma_de[4]_min	9	voiceProb_sma_de_min	5
pcm_fftMag_mfcc_sma_de[4]_range	9	voiceProb_sma_de_linregerrQ	5
pcm_fftMag_mfcc_sma_de[4]_stddev	9	voiceProb_sma_de_stddev	5
pcm_fftMag_mfcc_sma_de[5]_minPos	9	voiceProb_sma_de_skewness	5
pcm_fftMag_mfcc_sma_de[5]_linregerrQ	9	F0_sma_de_linregerrQ	5
pcm_fftMag_mfcc_sma_de[5]_stddev	9	F0_sma_de_stddev	5
pcm_fftMag_mfcc_sma_de[6]_linregc2	9	jitterLocal_sma_de_kurtosis	5
pcm_fftMag_mfcc_sma_de[11]_minPos	9	shimmerLocal_sma_de_maxPos	5
pcm_fftMag_mfcc_sma_de[11]_stddev	9	shimmerLocal_sma_de_linregerrQ	5
pcm_fftMag_mfcc_sma_de[12]_range	9	shimmerLocal_sma_de_stddev	5

➢ ANOVA分析

左上表格為使用ANOVA分析篩選之重要特徵值表格。將一階偏微分特徵值之累積出現次數由大至小排列，可看出最高之累計次數為10，而本研究取用累積出現次數大於等於6且小於等於10之一階偏微分特徵值進行後續建模。

➢ Spearman分析

左下表格為使用Spearman分析篩選之重要特徵值表格。將一階偏微分特徵值之累積出現次數由大至小排列，可看出最高之累計次數為11，而本研究取用累積出現次數大於等於6且小於等於11之一階偏微分特徵值進行後續建模。

模型成效比較

各方法之結果計算其平均準確率

準確率之平均值				
分析方法	特徵值使用			
	192d'	方法一	方法二	吳允彤(2022)
迴歸方法	22.01%	30.81%	28.30%	25.40%
分類方法	23.68%	28.46%	27.70%	27.74%

當使用全部一階偏微分特徵值(192d')進行建模時，其迴歸方法及分類方法之平均準確率分別為22.01%及23.68%，而使用方法一篩選後之特徵值進行建模，其迴歸方法與分類方法之平均準確率分別為30.81%和28.46%；使用方法二篩選後之特徵值在經過建模後，其迴歸方法與分類方法之平均準確率分別為28.3%和27.7%。兩種方法的成效皆高於使用全部一階偏微分特徵值，其中，又以方法一的迴歸方法成效最佳。

複迴歸分析

預測結果範例（閾值為3.9）

Actual value	Predicted value
5.0	3.061933
5.0	5.038622
2.0	5.386681
6.0	5.984404
4.0	3.961722
3.0	3.678023

第一次預測 Confusion Matrix	總個數	錯誤個數
	6	4

		實際	
		心情低	正常
預測	心情低	1	1
	正常	1	3

將受測者資料使用交叉驗證，並透過複迴歸分析進行預測後，計算每一次在不同閾值下的Hit rate、Miss rate、False alarm rate、Correct rejection rate四個數值，並計算十次結果的平均。我們將實際心情值3以下定義成心情低，心情值4以上定義成正常，而預測之心情值則是依照不同的閾值區分成心情低及正常兩類情緒。

複迴歸十次結果平均值

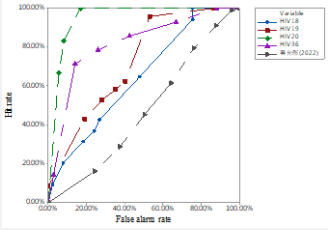
	閾值	3	3.3	3.7	3.8	3.9	4.1	4.4	4.7	5.2	5.4
HIV 18	Hit	9.26%	20.37%	31.48%	37.04%	42.59%	48.15%	64.81%	72.22%	94.44%	100.00%
	Miss	90.74%	79.63%	68.52%	62.96%	57.41%	51.85%	35.19%	27.78%	5.56%	0.00%
	False alarm	2.22%	7.78%	18.33%	23.89%	26.67%	35.00%	47.78%	58.89%	75.56%	75.56%
	Correct rejection	97.78%	92.22%	81.67%	76.11%	73.33%	65.00%	52.22%	41.11%	24.44%	24.44%

	閾值	3	3.2	3.5	3.6	4	4.1	4.6	4.8
HIV 19	Hit	0.00%	16.67%	29.17%	59.72%	80.56%	86.11%	95.83%	100.00%
	Miss	100.00%	83.33%	70.83%	40.28%	19.44%	13.89%	4.17%	0.00%
	False alarm	0.00%	1.85%	4.07%	14.10%	25.32%	33.10%	56.80%	63.31%
	Correct rejection	100.00%	98.15%	95.93%	85.90%	74.68%	66.90%	43.20%	36.69%

	閾值	3	3.4	3.6
HIV 20	Hit	66.67%	83.33%	100.00%
	Miss	33.33%	16.67%	0.00%
	False alarm	5.56%	8.33%	16.67%
	Correct rejection	94.44%	91.67%	83.33%

	閾值	3	3.3	4.3	4.6	4.7	5.9	6.4
HIV 36	Hit	0.00%	21.43%	28.57%	57.14%	57.14%	92.86%	100.00%
	Miss	100.00%	78.57%	71.43%	42.86%	42.86%	7.14%	0.00%
	False alarm	2.78%	2.78%	20.37%	25.93%	31.48%	91.67%	97.22%
	Correct rejection	97.22%	97.22%	79.63%	74.07%	68.52%	8.33%	2.78%

ROC曲線



當ROC曲線越貼近左上角表示成效越好，因此可以從左圖看出利用複迴歸分析所預測之四位均勻受測者的結果皆比吳允彤(2022)的研究較佳，且明顯看出受測者編號20的結果為最佳。

結論與討論

第一階段使用所有個別受測者在ANOVA和Spearman分析中出現累計次數大於六次之顯著特徵值放入模型中，準確率落在30%左右。而在第二階段針對個人去發展模型，同樣使用ANOVA和Spearman分析來挑選個人顯著特徵值，再加上利用複回歸分析以及套用品管中管制下限的概念，篩選不同的閾值，成效會較先前使用群體資料來預測個人資料的結果佳。在一樣是要找出80%的情緒低落狀況時，誤警能從原本75%降低至40%。然而使用複回歸分析來進行情緒預測仍會受到資料的限制，若是能夠收集充足且情緒分佈平均的資料，才可建構出合適的模型並準確地預測其情緒。