

巨量資料 專案報告

P002: IoT-based Virtual Health Determination

410978004 統計四 郭依璇
410978038 統計四 戴衣伶
410978051 統計四 林瑋珈
410978055 統計四 薛珮妤
410978058 統計四 李博業
411078018 統計三 侯言蓁
411078028 統計三 陳皓鈞



目錄

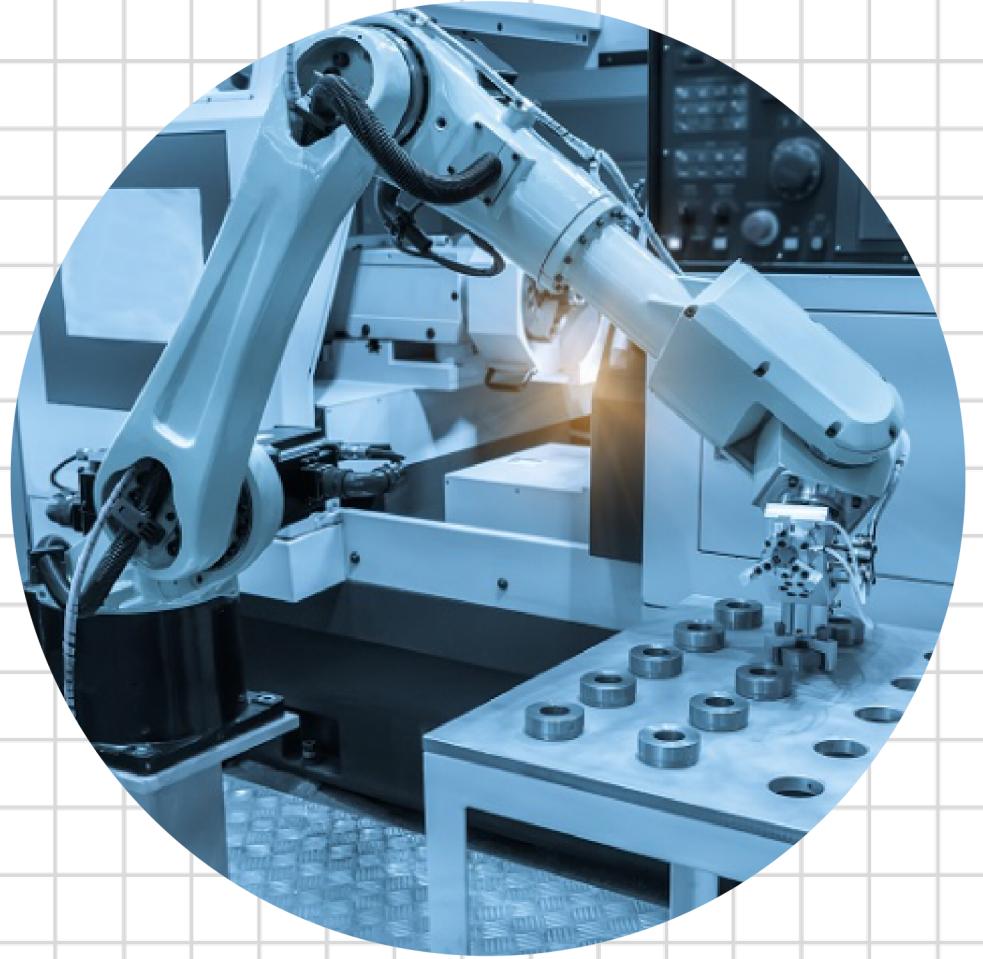
1.	專案回顧	2	5.	特徵擷取比較	35
2.	資料前處理	7	6.	建議決策方式	38
3.	快速傅立葉轉換.....	14	7.	輔助決策工具設計稿	40
4.	模型選擇	21	8.	Q&A	42





專案回顧

專案背景



機械手臂

- 組裝、搬運、焊接等



振動感測器

- 監測機手臂振動情況

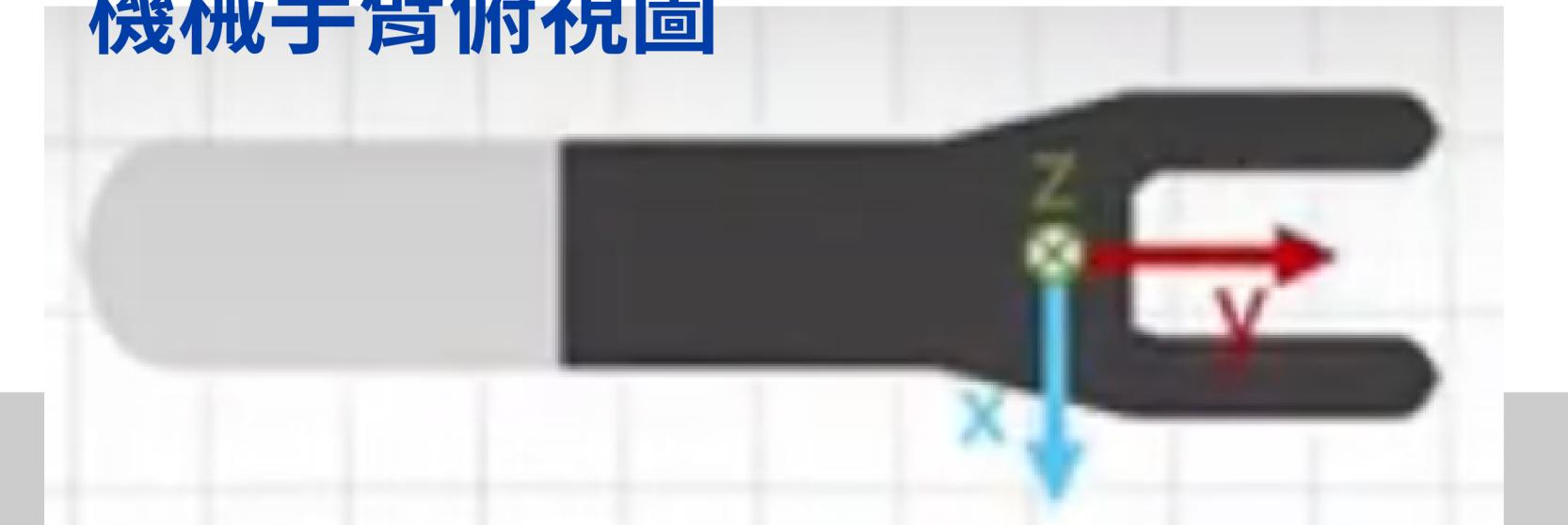


監測、分析

- 對機械手臂健康狀態的及時診斷與預測

機械手臂作動

機械手臂俯視圖



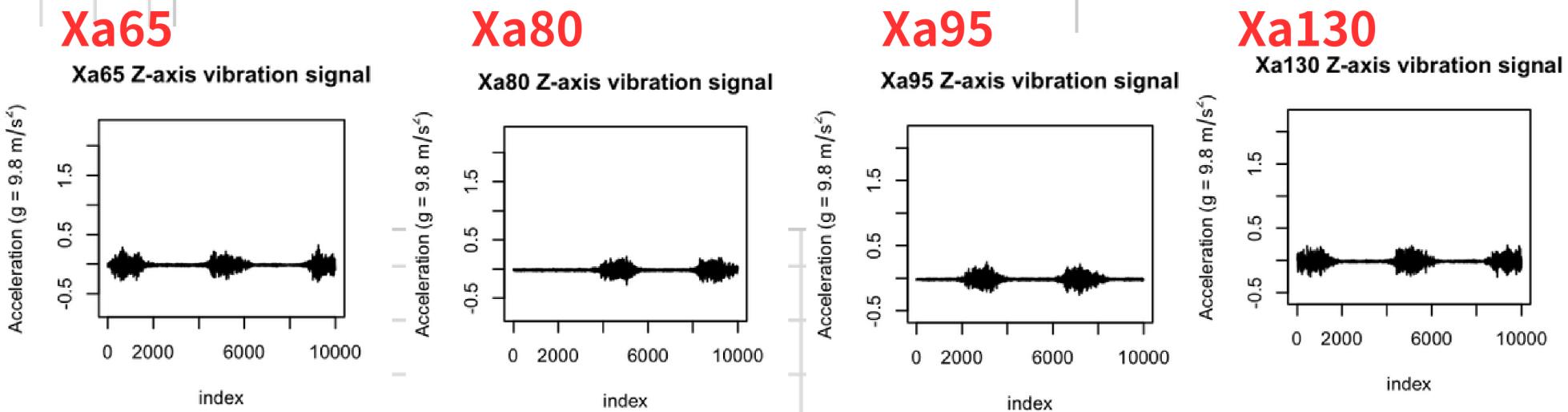
資料說明

資料：安裝於機械手臂的三軸加速規所感測到的 x、y、z 三軸向振動訊號。

- 共有四顆三軸加速規，分別安裝於：
 - * Xa - 位於機械手臂**水平作動方向傳動軸的馬達側**
 - * Xb - 位於機械手臂**水平作動方向傳動軸的惰輪側**
 - * Ya - 位於機械手臂**垂直作動方向傳動軸的馬達側**
 - * Yb - 位於機械手臂**垂直作動方向傳動軸的惰輪側**
- 再對相異方向的傳動軸施加不同程度的負荷，分別如下：
 - * 水平方向：65、80、95、130，其中**正常負荷約於 80 左右**
 - * 垂直方向：220、260、300、380，其中**正常負荷約於 260 左右**
- 採集設置：固定張力下，機械手臂以固定速度、模式作動約 5 - 6 分鐘，
 振動訊號每隔 5 秒儲存成單一檔案

上次報告 Q&A 回顧

文獻資料跟我們的
有什麼差別？



不同施壓力震動基準相同，文獻是明顯不同的

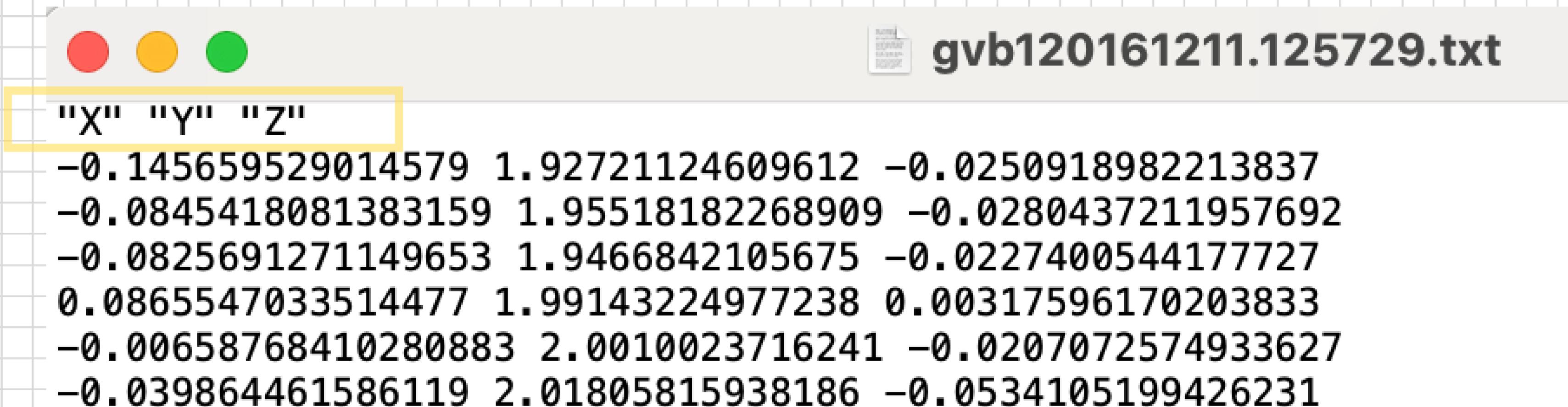
處理完資料後，
嫡跟傅立葉是會套
同個模型分類器嗎？

建模型前是否先檢定特徵？



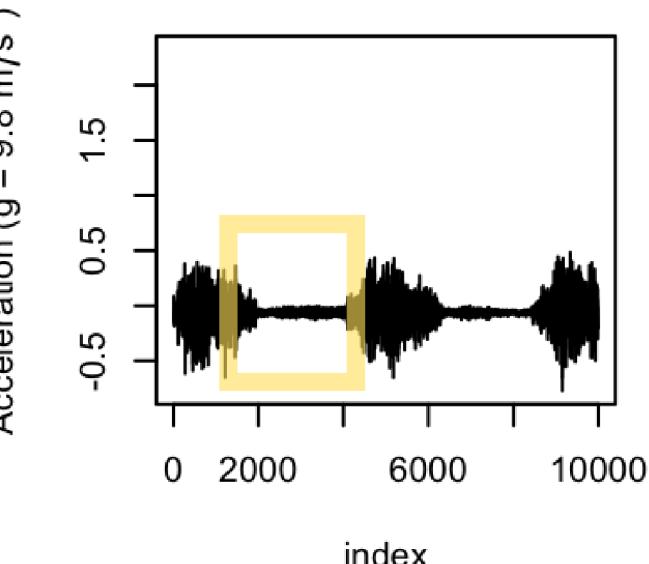
資料前處理

Xa65 raw data (.txt檔)



gvb120161211.125729.txt

X	Y	Z
-0.145659529014579	1.92721124609612	-0.0250918982213837
-0.0845418081383159	1.95518182268909	-0.0280437211957692
-0.0825691271149653	1.9466842105675	-0.0227400544177727
0.0865547033514477	1.99143224977238	0.00317596170203833
-0.00658768410280883	2.0010023716241	-0.0207072574933627
-0.039864461586119	2.01805815938186	-0.0534105199426231

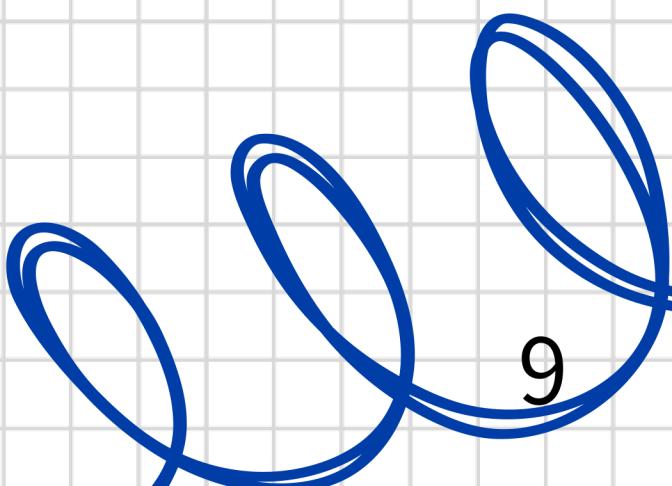


回顧：找出波的頭尾

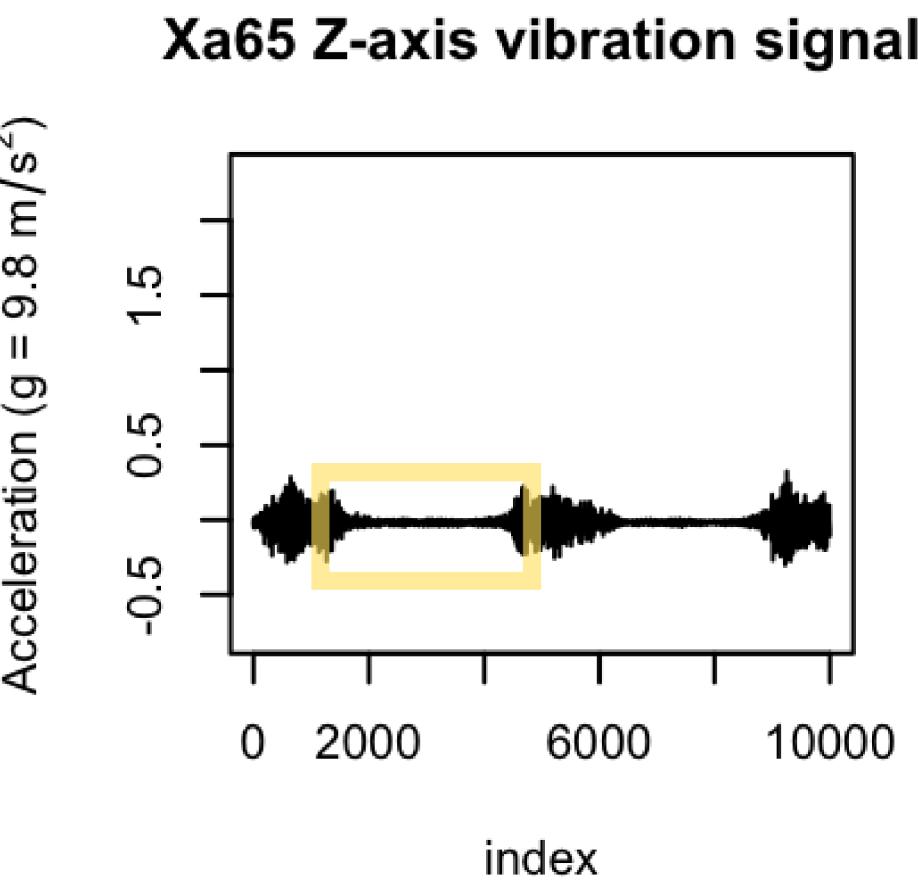
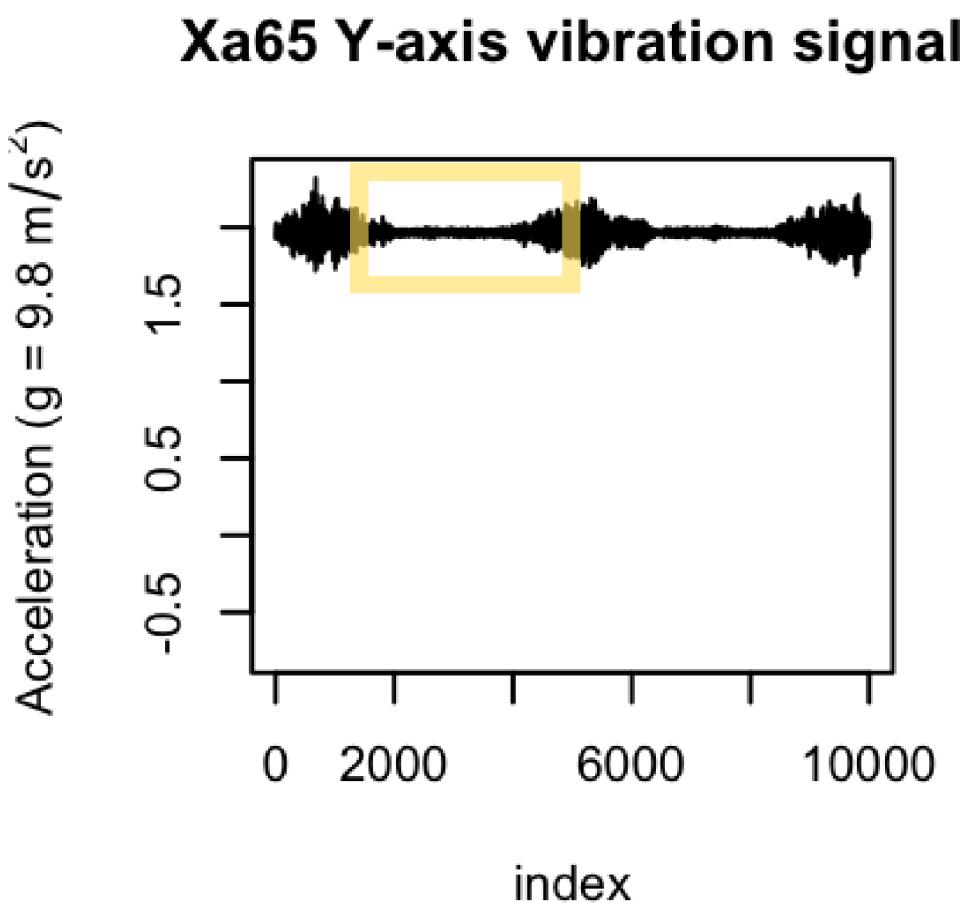
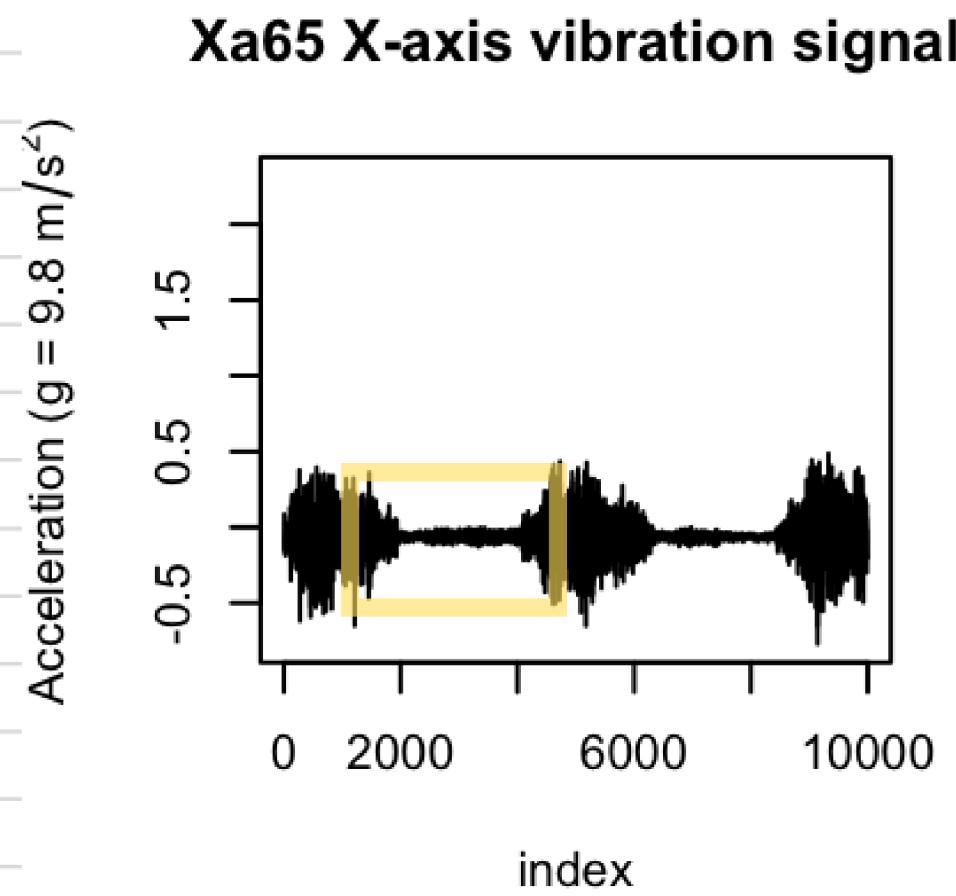
1. 觀察非振動期訊號的 “各軸” range 分佈在哪個區間
2. 找出非振動期訊號的頭、尾，就是波的尾、頭
3. 確保每個波的頭尾都找到後，只匯出在頭尾內的資料

此法可確保波內之值不會被刪掉，並且可幫助後續傅立葉轉換

=> 頭尾的index決定，改為以X軸找到的頭尾為主



Xa65於X.Y.Z軸之振動訊號圖



range明顯不同

進出不同軸在振動期間的資料

gvb000001_X_ht

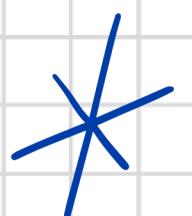
Segment_Head	Segment_Tail	Index	Data
1	1903	1	-0.17139538261847
1	1903	2	-0.23974250955062

gvb000001_Y_ht

Segment_Head	Segment_Tail	Index	Data
1	1912	1	1.92030137031769
1	1912	2	1.88935298080625

gvb000001_Z_ht

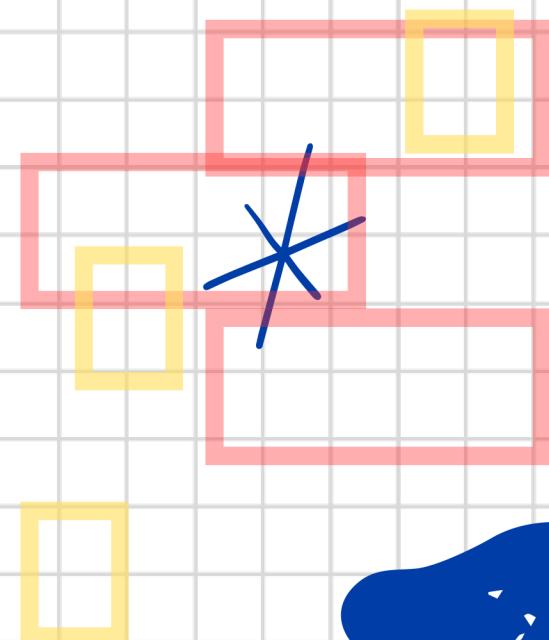
Segment_Head	Segment_Tail	Index	Data
1	1533	1	-0.0377485709666376
1	1533	2	-0.0811374044657039



進出不同軸在振動期間的資料

gvb000001_X_ht

Segment_Head	Segment_Tail	Index	Data
1	1903	1	-0.17139538261847
1	1903	2	-0.23974250955062



gvb000001_Y_ht

Segment_Head	Segment_Tail	Index	Data
1	1912	1	1.92030137031769
1	1912	2	1.88935298080625



gvb000001_Z_ht

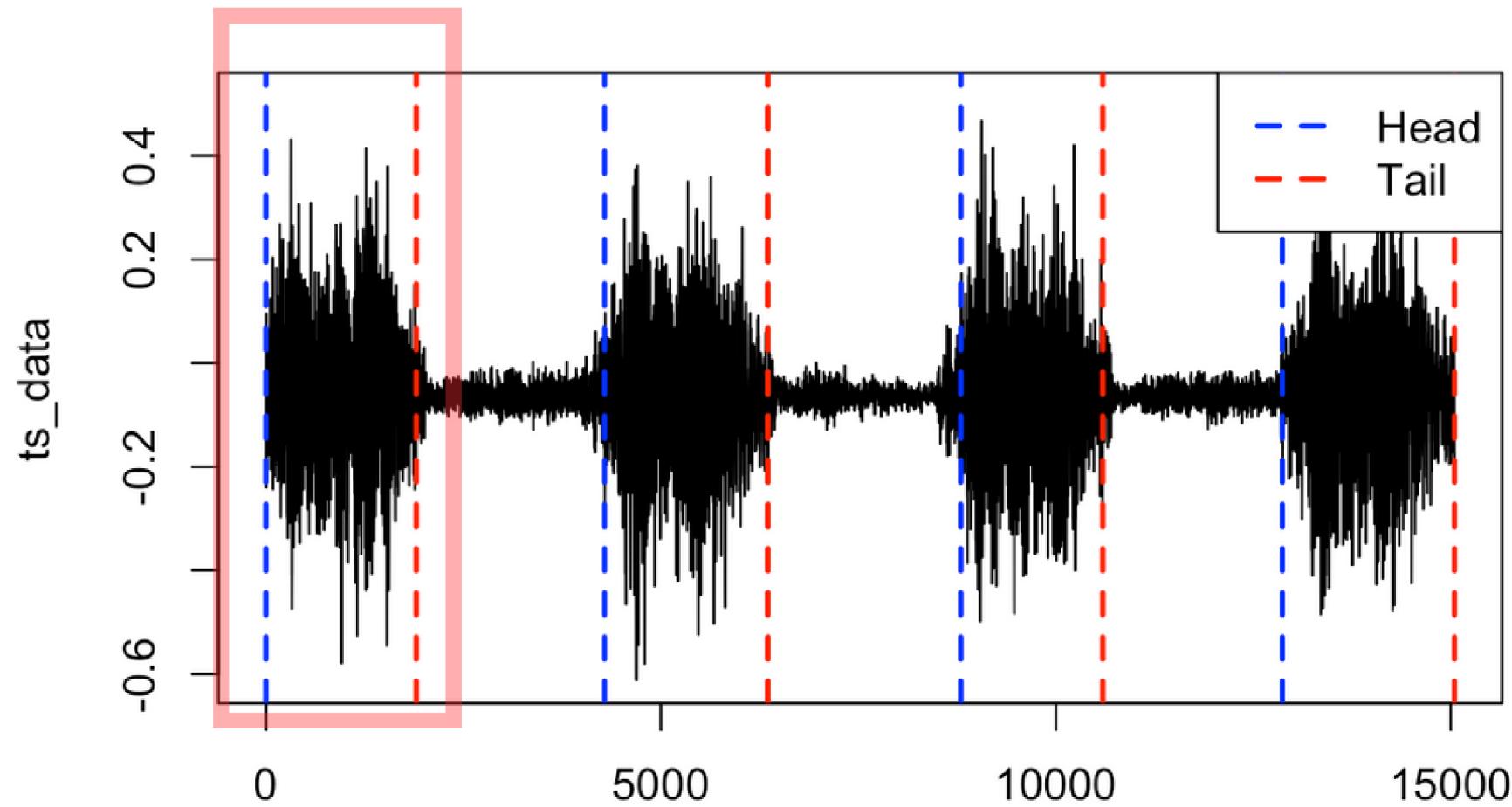
Segment_Head	Segment_Tail	Index	Data
1	1533	1	-0.0377485709666376
1	1533	2	-0.0811374044657039

找波的頭尾--以X軸為主

gvb000001_X_ht

Segment_Head	Segment_Tail	Index	Data
1	1903	1	-0.17139538261847
1	1903	2	-0.23974250955062

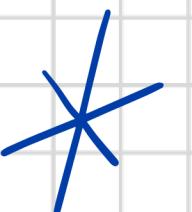
gvb000001.txt X-axis



匯出皆在振動期間的資料

gvb000001_ht

Segment_Head	Segment_Tail	Index	Xaxis	Yaxis	Zaxis
1	1903	1	-0.17139538261847	1.94027547943536	-0.0158814856771068
1	1903	2	-0.23974250955062	1.88446674750283	-0.0260903061480557
1	1903	3	-0.22841744103319	1.8851613387312	-0.0471376750625477
1	1903	4	-0.110896315477446	1.93232799168252	-0.0378148184417492
1	1903	5	0.0104970030885436	1.98218778405081	-0.0207927941155563
1	1903	6	0.0952942072288055	2.03958797375027	-0.0226872748073539
1	1903	7	0.0265493187502443	2.03621273426513	-0.0247143771592535
1	1903	8	-0.0553261931121386	2.02126391556015	-0.0202879986011843
1	1903	9	-0.0555001017690572	2.02006911624497	-0.0202466503185839
1	1903	10	-0.176590559583117	1.93714681160767	-0.00756931381864525





快速傳立葉轉換



欄位介紹

欄位名稱	欄位定義
FFT_x,y,z	複數形式 $X[k] = A[k] + jB[k]$ 實部：餘弦成分；虛部：正弦成分
Frequency	$f_k = \frac{k \cdot f_s}{N}$
FFT_x,y,z_Amplitude	$ X[k] = \sqrt{A[k]^2 + B[k]^2}$

資料呈現

Index	Xaxis	Yaxis	Zaxis	FFT _x	FFT _y
1	-0.135806989368321	1.89515804765186	-0.0148972179044471	(-131.31624418897928+0j)	(3978.967010227832+0j)
2	-0.213469021086208	1.91450482511597	-0.014869576385514	(2.5903906956645764-1.4994116554472594j)	(0.5442765484775022+0.04411071564599339j)
3	-0.132001950370146	1.93689920205785	-0.0134229907971671	(0.49801805739196003-1.3439324398161432j)	(0.30705068918183775+0.231654211527863j)
4	-0.0417560639949536	1.98273929952102	-0.0334715810869589	(-0.10444691174620857+0.9187915336246351j)	(-0.09112895320835657-0.9719785905388034j)
5	-0.0310658649649734	1.99595970400162	-0.0295356211320834	(2.995259928854592+0.03576004928337562j)	(0.6885658296467566+0.44902670481163454j)
6	-0.0139184364330957	2.01848543592502	-0.0031692194563537	(-4.435741479907254-0.049248766788606235j)	(-1.2140265623346278-1.0325755726640642j)
7	0.0075533611244761	2.02259130857946	-0.015907430814932	(-0.6344104361131415+3.476472949914047j)	(0.934185090900877+0.5768684447443467j)
8	-0.0163240641257373	1.99674203105104	-0.0198178667030123	(0.4231324475245368-0.4998656226148612j)	(0.38411893243671824+0.8262495492071826j)
9	-0.124794309272709	1.96449636845713	0.004633137382024	(-0.7769517562131297+0.8321299442857604j)	(-0.9466814598862472-0.9969536083545374j)

FFT _z	Frequency	FFT _{x_Amplitude}	FFT _{y_Amplitude}	FFT _{z_Amplitude}
(-37.592627334675775+0j)	0.0	131.31624418897900	3978.967010227830	37.592627334675800
(0.7398112122004701+0.6046859466646202j)	0.0004938271604938270	2.9930518319395500	0.5460610922391230	0.9554923986046240
(-0.5392264851813333-0.9338420010025502j)	0.0009876543209876540	1.4332398223181700	0.38463463110546500	1.078344140410410
(0.3758724832219209+0.7019694079385534j)	0.0014814814814814800	0.9247091648911180	0.9762411928302610	0.7962670238839590
(0.7301064365611484+0.5990595476371148j)	0.0019753086419753100	2.995473388719510	0.822038857591976	0.9444192661753590
(-0.10693788640533303-0.43574467919572984j)	0.0024691358024691400	4.436014868956140	1.593760586573950	0.44867486780098300
(-0.3277495859895381+0.5133271737627823j)	0.0029629629629629600	3.533884657559350	1.0979430707478600	0.6090357776351070
(-0.09545480331751666-0.3714050969107684j)	0.0034567901234567900	0.6549096951643390	0.9111726904497290	0.3834753779419
(0.7804751769962959+1.2634445486173371j)	0.003950617283950620	1.138461363270480	1.3748171819203800	1.4850702439070900

特徵擷取 - 波形基本特徵

特徵類別

- 變異數、分位數、平均數、中位數 (`var_`, `q1_`, `q3_`, `mean_`, `meadian_`)：敘述性統計
- 主頻率 (`dominant_frequency_`)：振幅最大的頻率
- 峰值數量 (`number_of_peaks_`)：峰值的數量
- 峰值振幅 (`peak_amplitude_`)：所有峰值中的最大振幅
- 頻譜能量 (`spectral_energy_`)：振幅的平方和，也可看作整段波的能量總和
- 頻譜熵 (`spectral_entropy_`)：頻譜的複雜度與不規則性
- 頻率帶寬 (`bandwidth_`)：主要能量的頻率範圍
- 頻率中心 (`spectral_centroid_`)：訊號能量分佈的重心
- 頻譜尖銳度 (`spectral_kurtosis_`)：測量頻譜形狀的峰度

資料表模樣

var_x	median_x	q1_x	mean_x	q3_x	dominant_frequency_x	number_of_peaks_x
5778.648326	9.900153	5.750465	15.441124	15.214783	0.0	0
5981.544065	9.186477	5.865434	14.473980	13.670499	0.0	0
5392.452547	9.289726	5.845974	14.713276	13.572046	0.0	0
6434.370071	9.213310	6.050526	14.745507	14.399582	0.0	0
6227.298020	10.137979	6.323567	15.498301	15.282176	0.0	0

peak_amplitude_x	spectral_energy_x	spectral_entropy_x	bandwidth_x	spectral_centroid_x	spectral_kurtosis_x
1673.033105	2.912265e+06	7.405788	0.0	0.187943	0.965094
1765.808187	3.219341e+06	7.427332	0.0	0.175950	0.972398
1577.802187	2.585718e+06	7.300492	0.0	0.181211	0.965968
1884.528535	3.665142e+06	7.514677	0.0	0.181756	0.972415
1810.577678	3.401903e+06	7.535032	0.0	0.188921	0.967760

特徵 Fisher 檢定

檢定不同位置下特徵是否能夠區分機械手臂的健康度

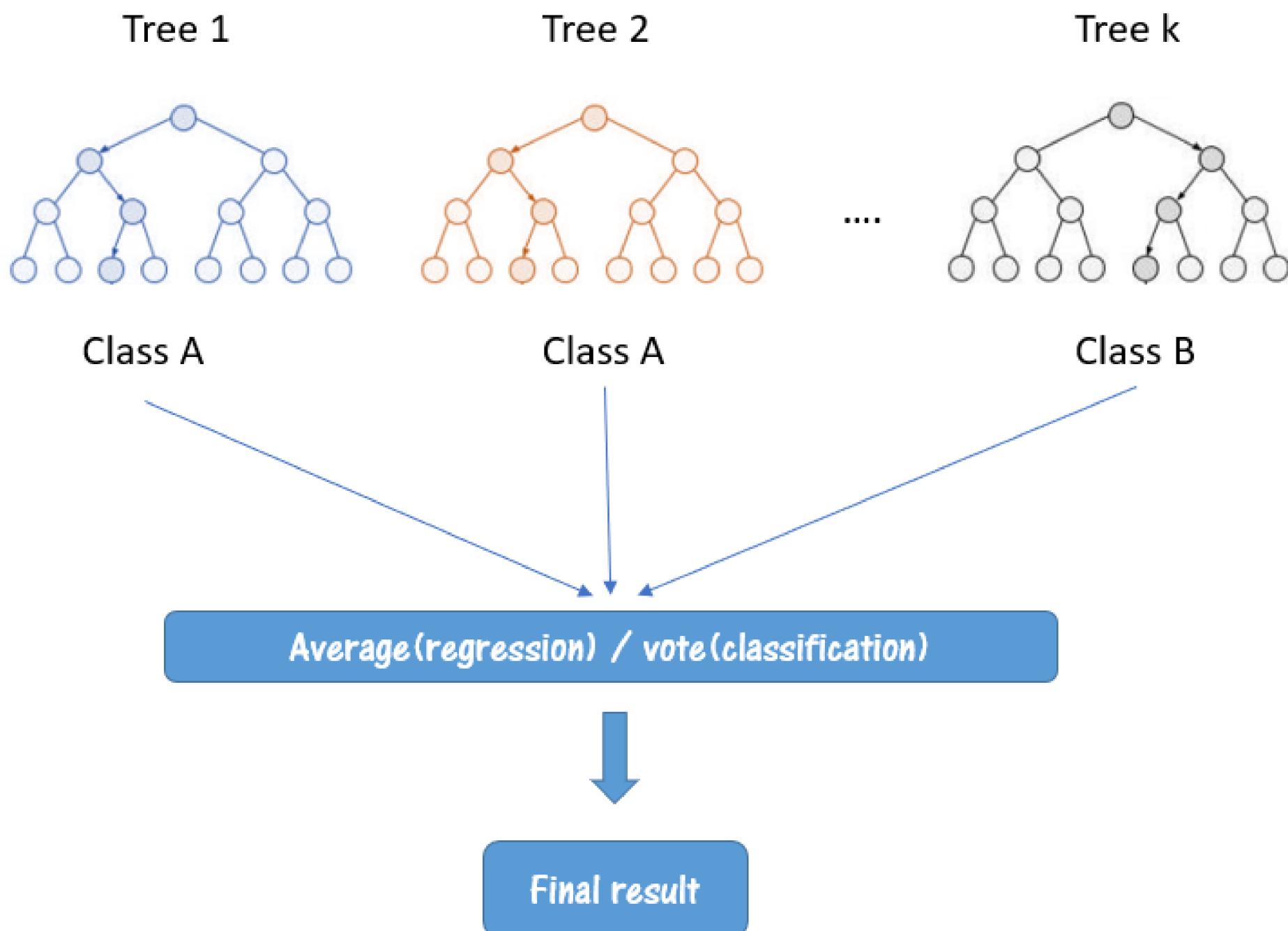
- X_a 、 Y_b 的變數對區分健康度的能力較弱
- X_b 、 Y_a 的變數對區分健康度的能力較強

特徵	X_a	X_b	Y_a	Y_b
變異數		√	√	
分位數		√	√	
平均數		√	√	
中位數		√	√	
主頻率			√	
峰值數量				√
峰值振幅		√	√	
頻譜能量		√	√	
頻譜熵		√	√	
頻率帶寬		√	√	√
頻率中心		√	√	
頻譜尖銳度		√	√	√



模型選擇

隨機森林



- 捕捉不同變數的共同交互結果
 - 有效率地試驗 X, Y, Z軸的交互效果
- 變數的隨機抽取
 - 增加模型多樣性(減少過擬合)
 - 減少變數共線性的影響

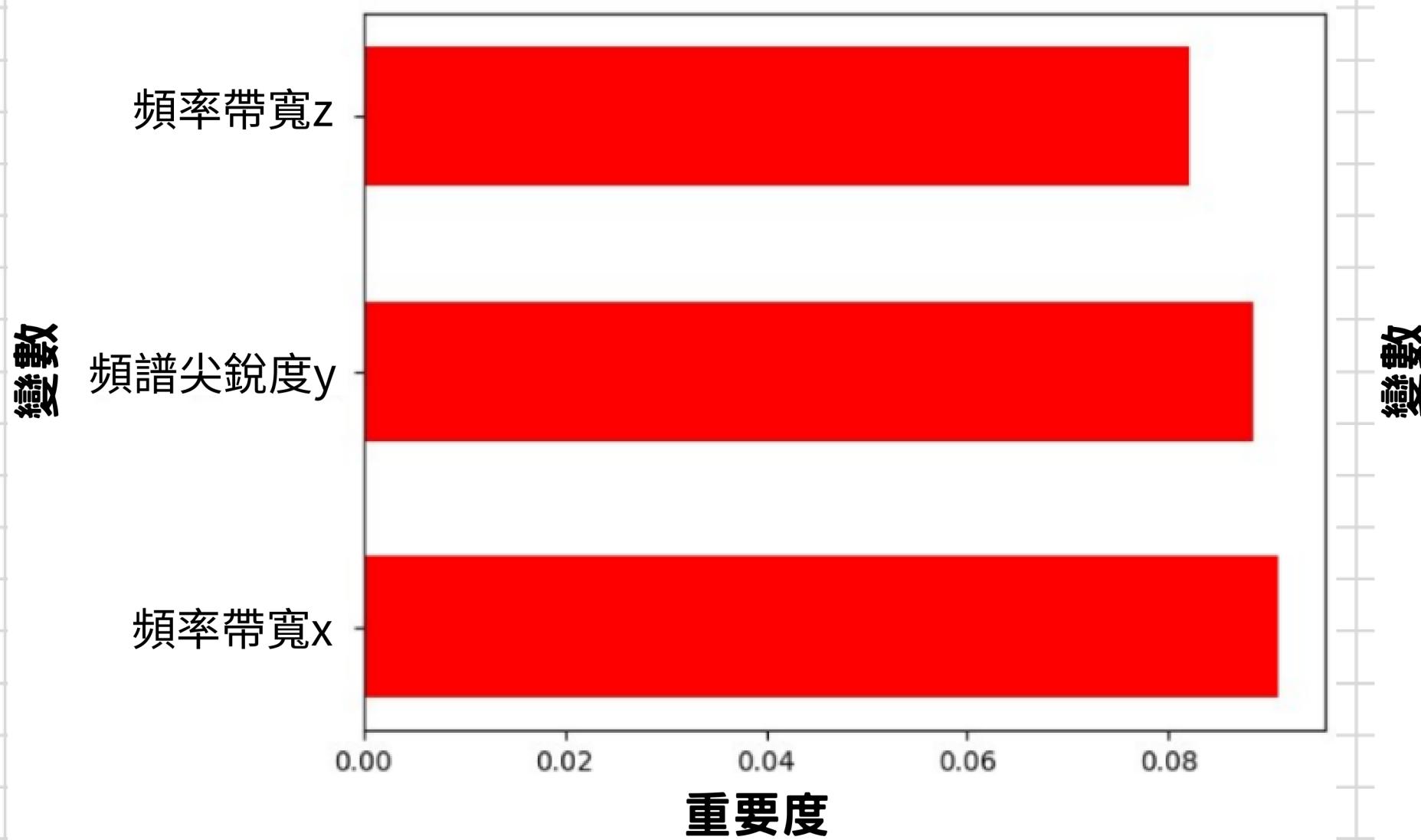
Model 1 - All Variables

隨機森林分類結果

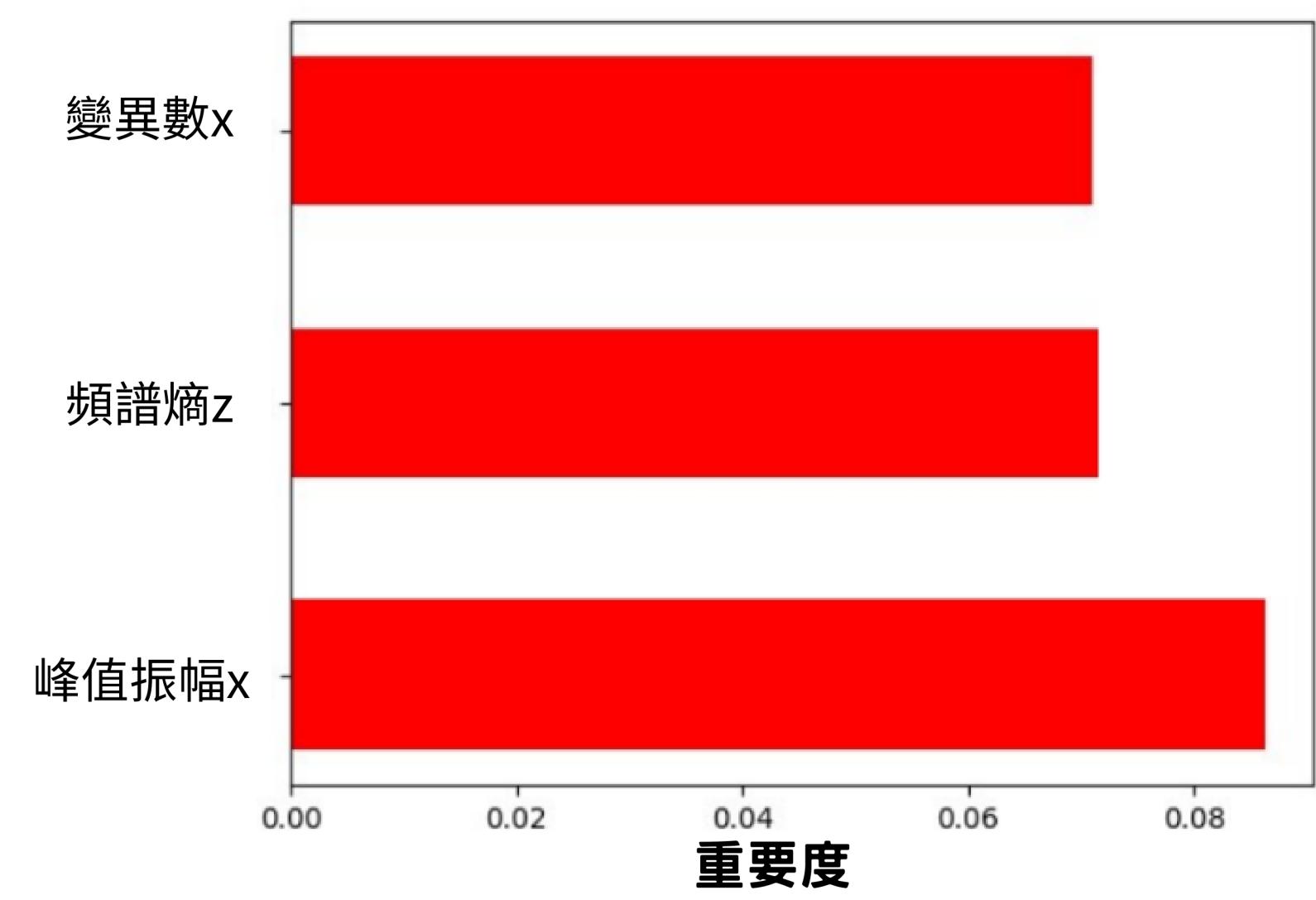
位置	模型準確率
水平作動馬達側(Xa)	0.8813
水平作動惰輪側(Xb)	0.9281
垂直作動馬達側(Ya)	0.9918
垂直作動惰輪側(Yb)	0.9836

隨機森林特徵重要度

水平作動馬達側(Xa)的變數前三重要度

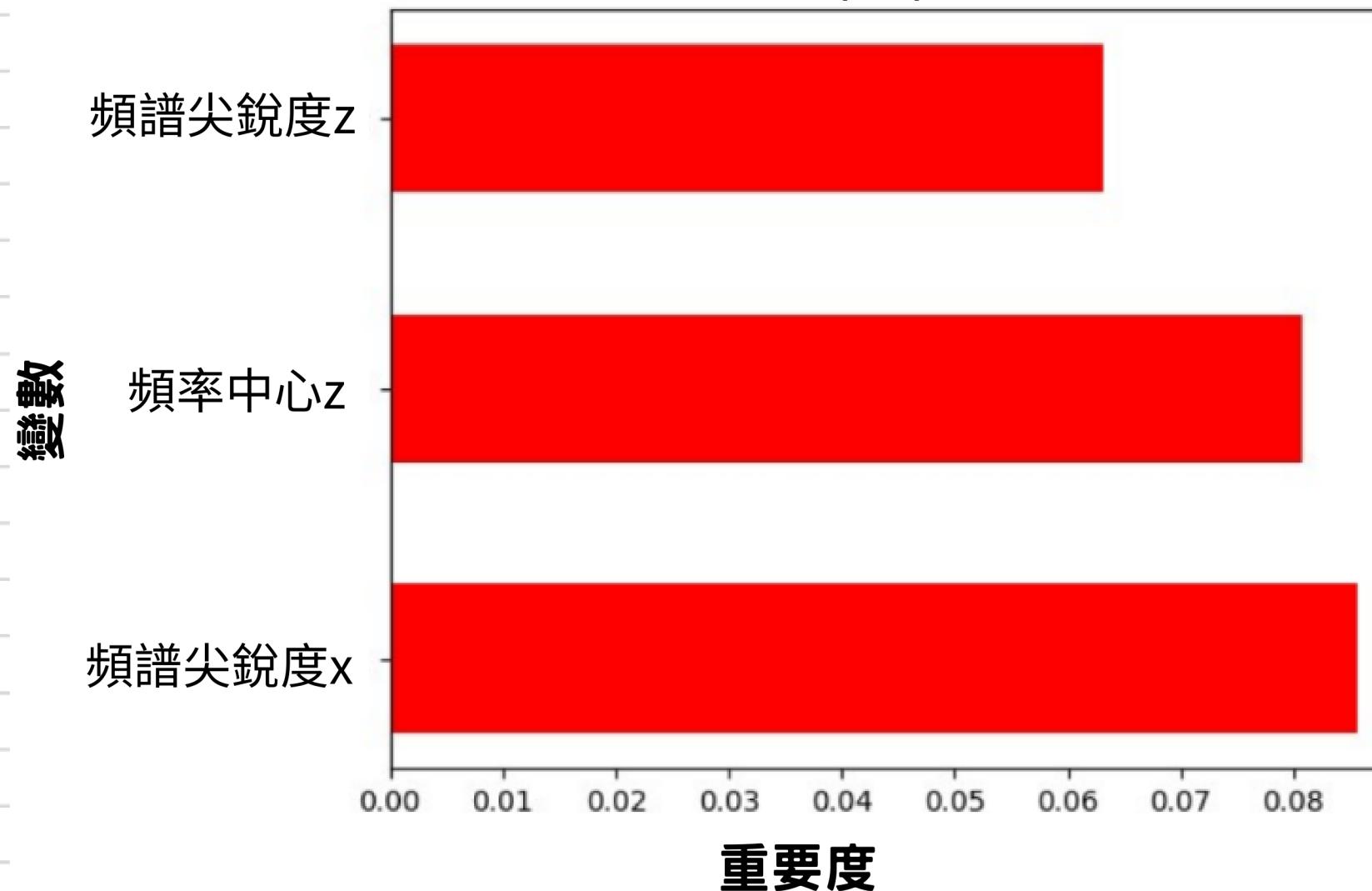


水平作動惰輪側(Xb)的變數前三重要度

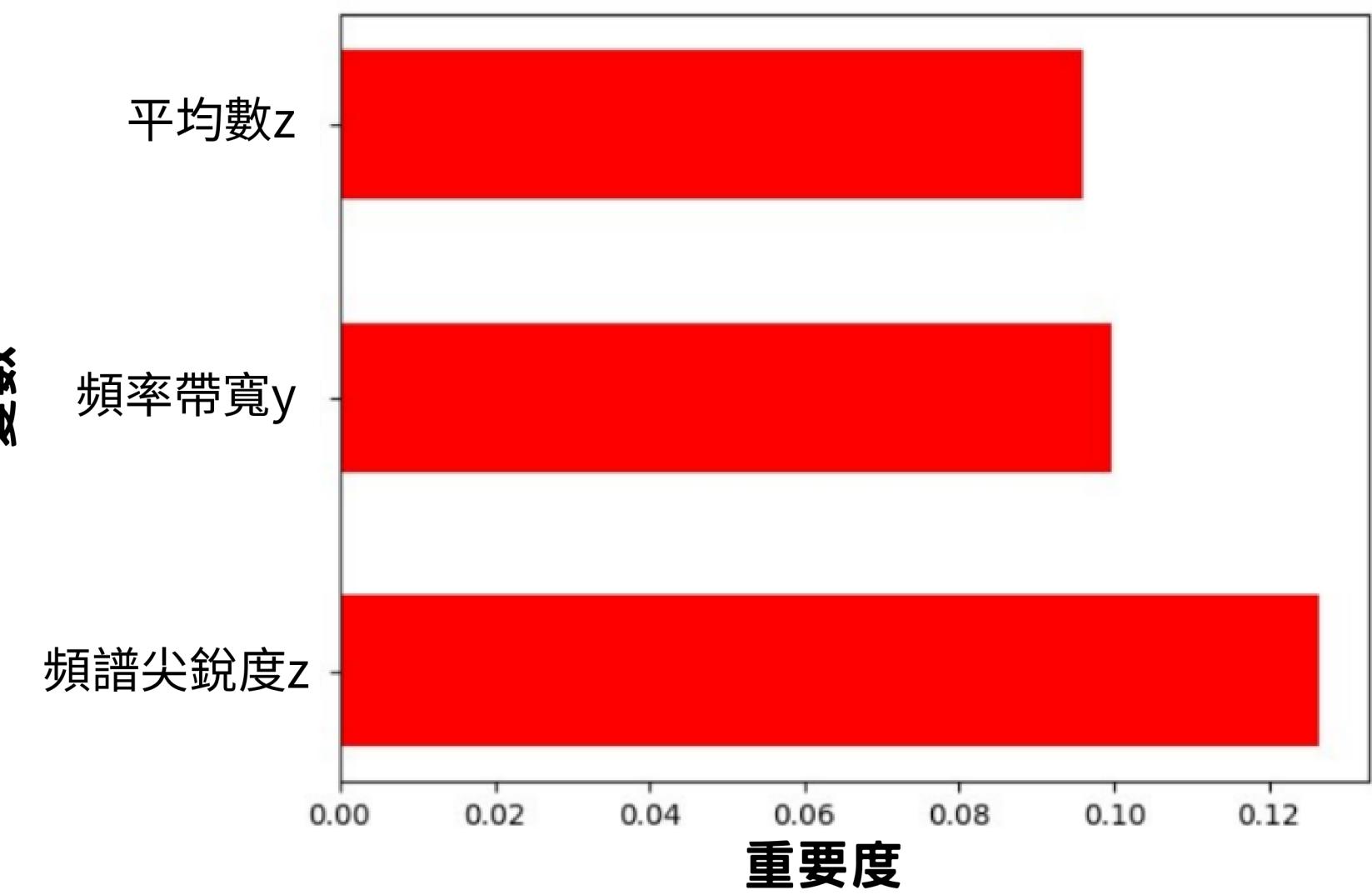


隨機森林特徵重要度

垂直作動馬達側(Ya)的變數前三重要度



垂直作動惰輪側(Yb)的變數前三重要度



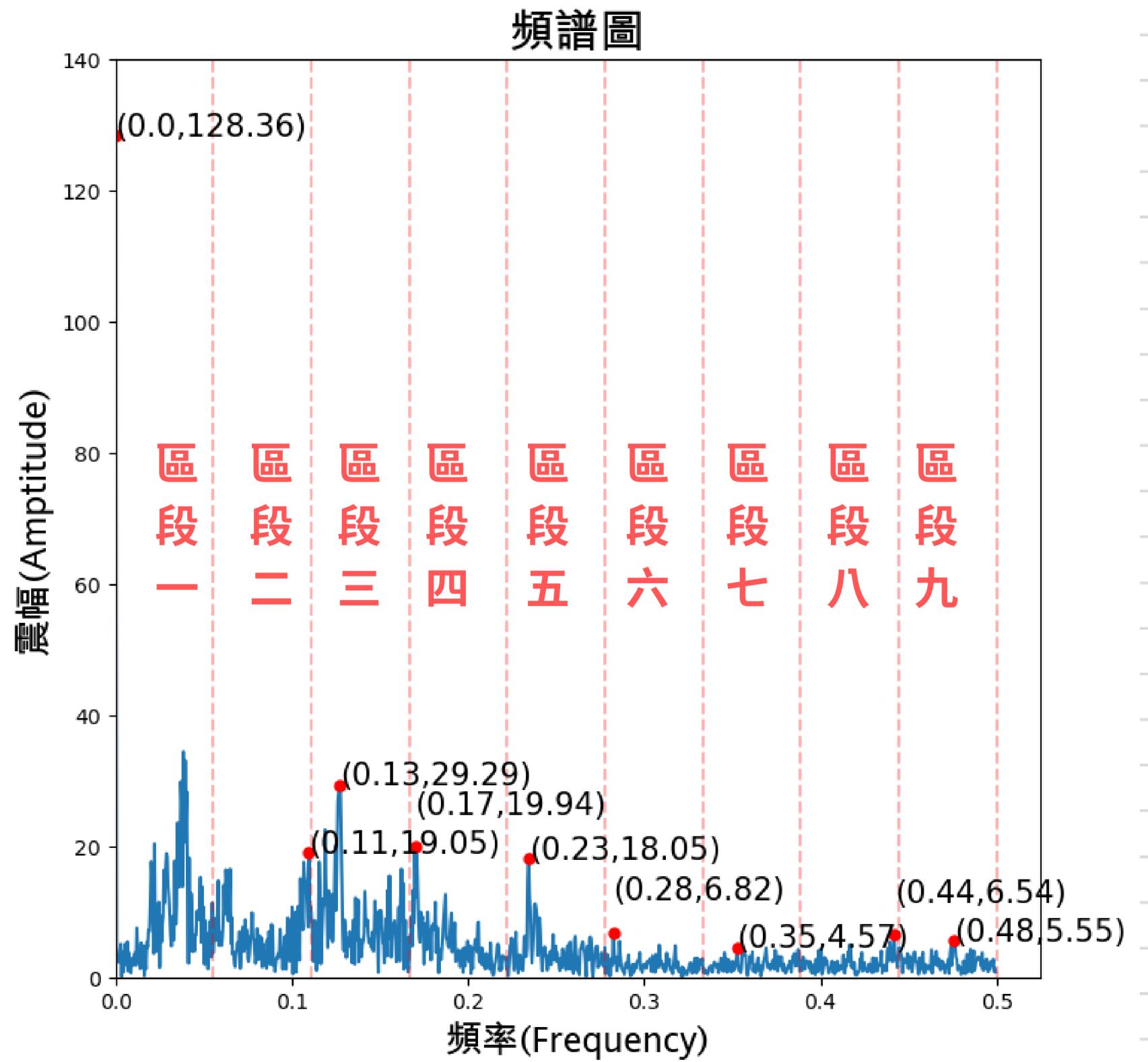
Model2 - Peak_Amplitude

Why 嘗試分段擷取極大值

- 振幅極大值為較原始的頻譜特徵
- 原特徵變數需要做參數選擇
(例如：帶寬範圍選擇)
- 往異常值方向思考，鎖定尋找離群值



分段擷取極大值示意圖



- 以Xa130X軸第一個波的頻譜圖為例
- 切片數 : 9
- 每一區塊取其極大振幅值(紅點)共九點

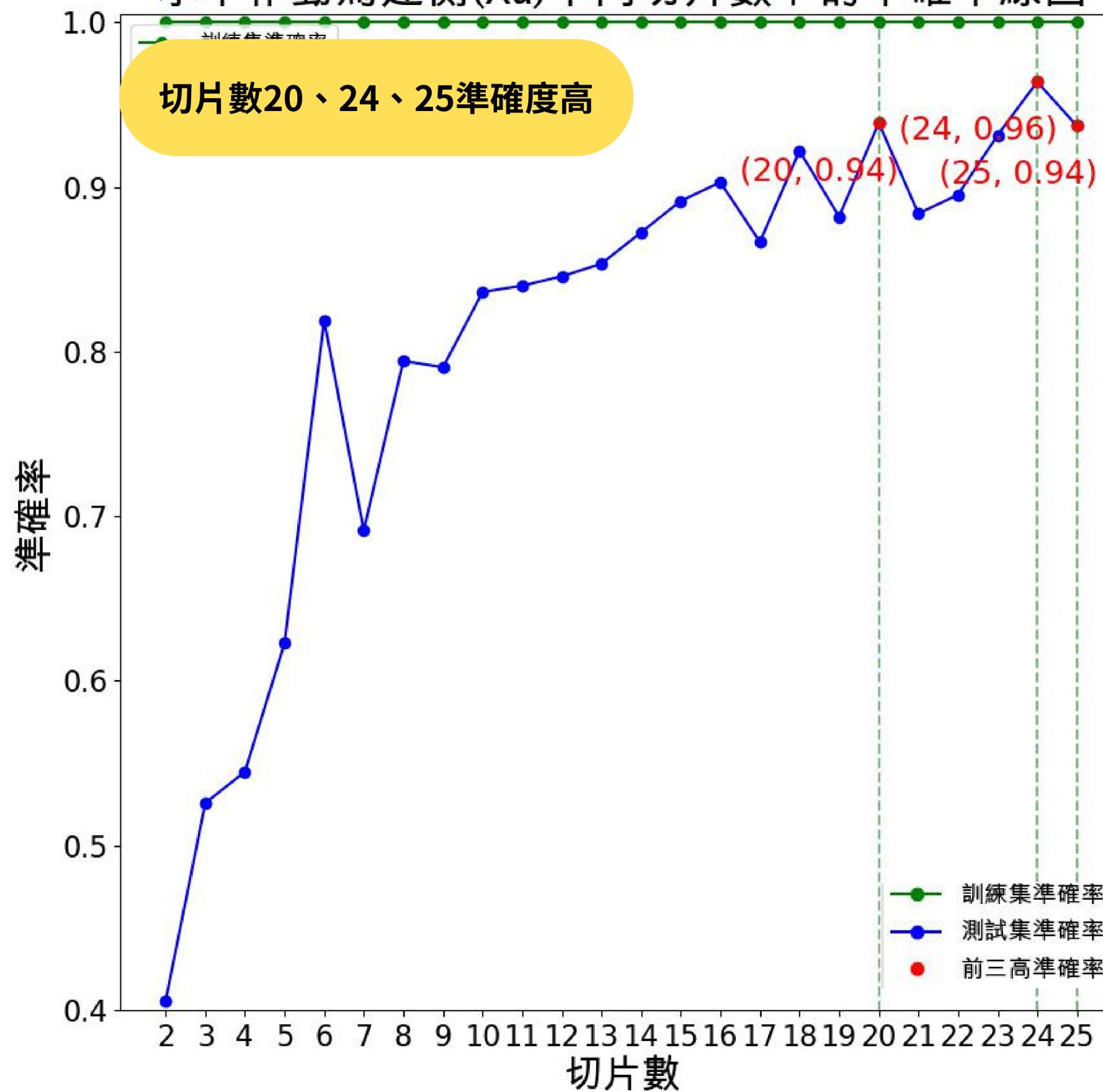


資料表模樣

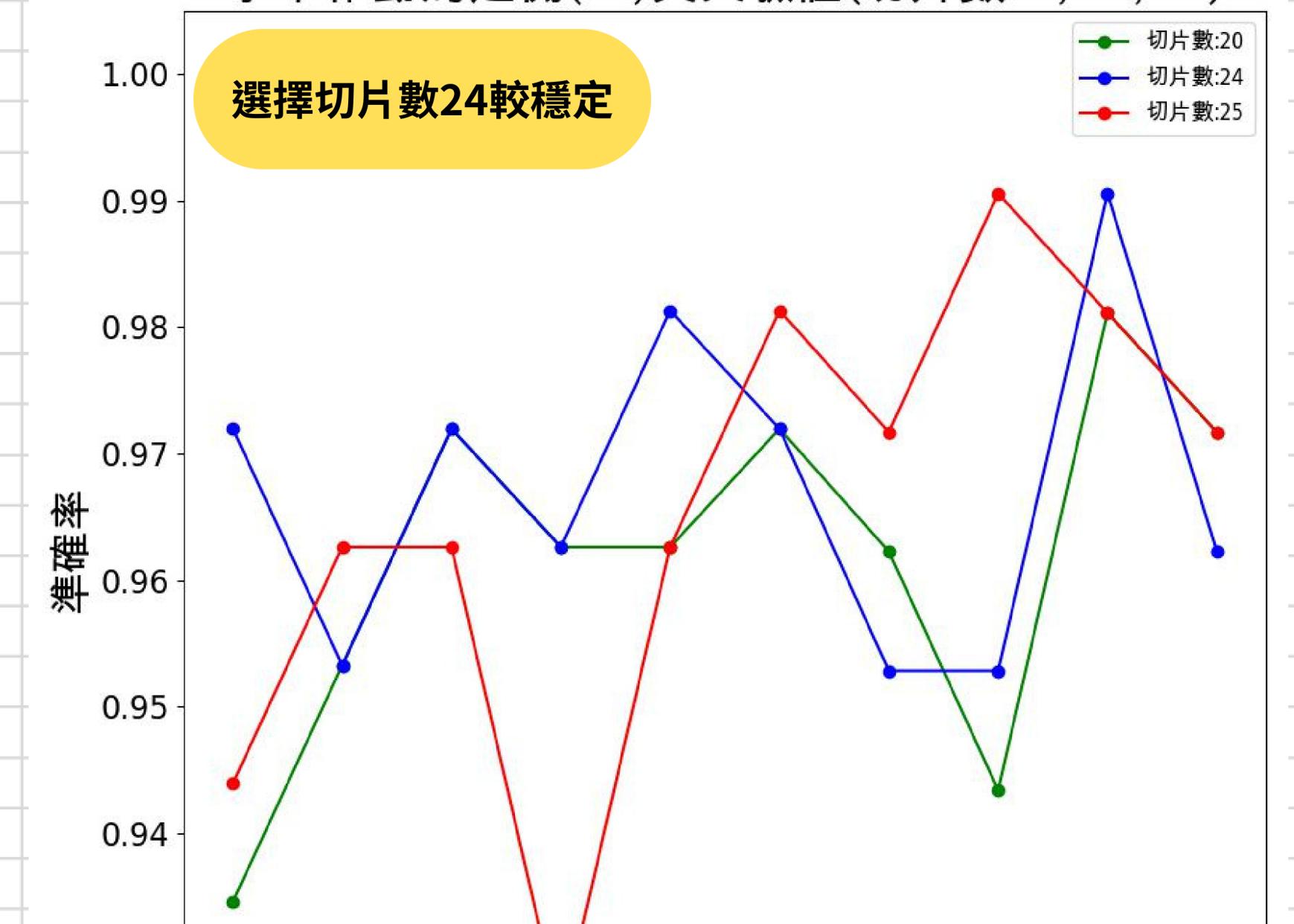
以水平作動馬達側(Xa)、切片數9為例

此波頭 Index	此波尾 Index	X軸-區段一	Y軸-區段一	Z軸-區段一	施壓 力種 類	X軸-區段二	Y軸-區段二	Z軸-區段二	X軸-區段三	...	Z軸-區段六	X軸-區段七	Y軸-區段七	Z軸-區段七	X軸-區段八	Y軸-區段八	Z軸-區段八	X軸-區段九	Y軸-區段九	Z軸-區段九
1	2078	128.357992	4083.127587	36.099974	Xa130	19.045750	14.166618	24.695610	29.293866	...	2.626262	4.569690	2.585464	2.276482	6.541043	4.089064	1.928640	5.552386	3.614318	2.201766
4544	6481	121.941518	3808.073081	33.078000	Xa130	22.251669	14.890844	11.624025	28.918423	...	2.159985	3.679214	2.256582	2.044124	4.350336	2.907558	1.879337	5.399984	2.286200	1.655007
8804	10760	124.393957	3846.764284	34.911610	Xa130	21.301692	15.989202	14.895162	28.040685	...	2.373301	4.212958	1.945391	2.036227	5.331618	3.091522	2.107222	5.276813	3.399689	1.956968
13094	15170	128.994947	4082.883088	37.294364	Xa130	15.290998	14.416036	17.967302	35.133458	...	2.343268	3.607413	2.282151	1.994728	4.559290	2.539331	1.571029	5.844351	2.654912	1.457851
17359	19418	128.390287	4051.190050	36.668222	Xa130	26.851615	12.316387	25.209910	32.410869	...	2.838323	3.770598	3.518398	2.070583	4.481467	2.291585	2.510453	5.588660	2.445783	2.222841
...	
1162004	1163830	115.424328	3588.981654	31.916747	Xa95	20.141984	14.144668	17.382877	25.162152	...	1.924958	3.752482	3.351303	1.903265	4.090498	2.420726	1.820957	5.394812	2.159510	1.716154
1166122	1168303	139.745859	4287.414459	36.515553	Xa95	17.841407	11.189808	23.785653	19.740897	...	3.087204	4.209344	2.340858	1.666500	5.423210	2.946645	2.374184	4.901717	2.058030	1.610647
1170755	1172536	110.915780	3503.886184	30.745418	Xa95	18.596026	13.839569	15.658039	21.673802	...	1.902451	3.396531	2.366184	1.575459	5.533193	2.969327	1.682535	4.251967	2.383180	1.629175
1174811	1176991	139.317341	4286.819476	38.242303	Xa95	14.163586	9.395683	17.898269	24.944048	...	2.675327	3.988610	2.243270	2.118932	5.103159	2.739998	1.975428	5.673748	2.087456	1.894744
1179466	1181228	116.107926	3465.299783	31.688821	Xa95	21.851793	17.169592	20.096285	22.535663	...	1.735784	3.177856	2.283248	1.580838	4.137297	2.420999	1.555832	4.088187	2.446283	1.422006

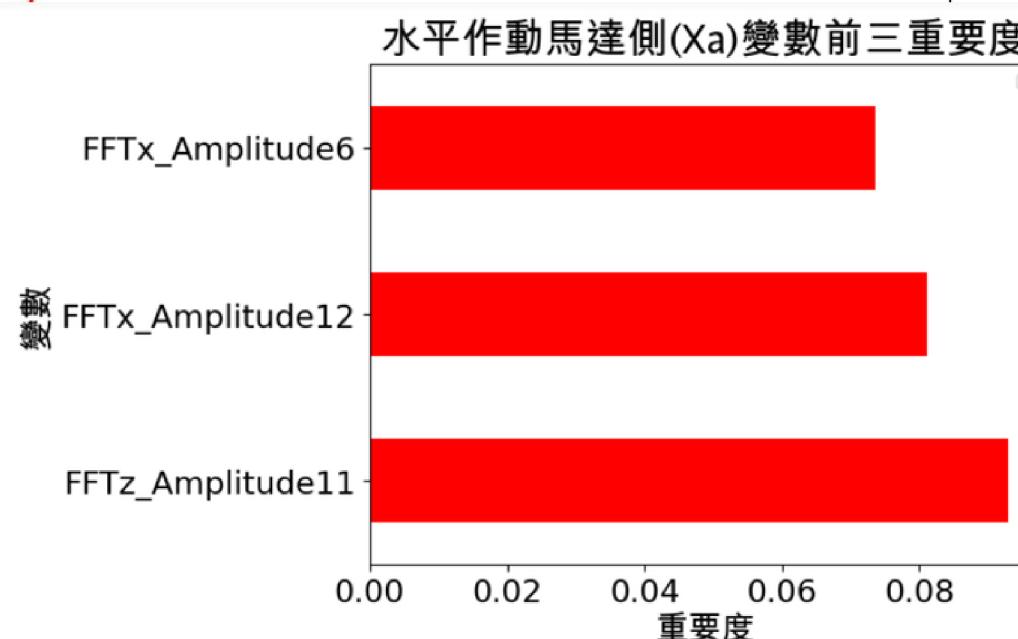
水平作動馬達側(Xa)不同切片數下的準確率線圖



水平作動馬達側(Xa)交叉驗證(切片數20, 24, 25)



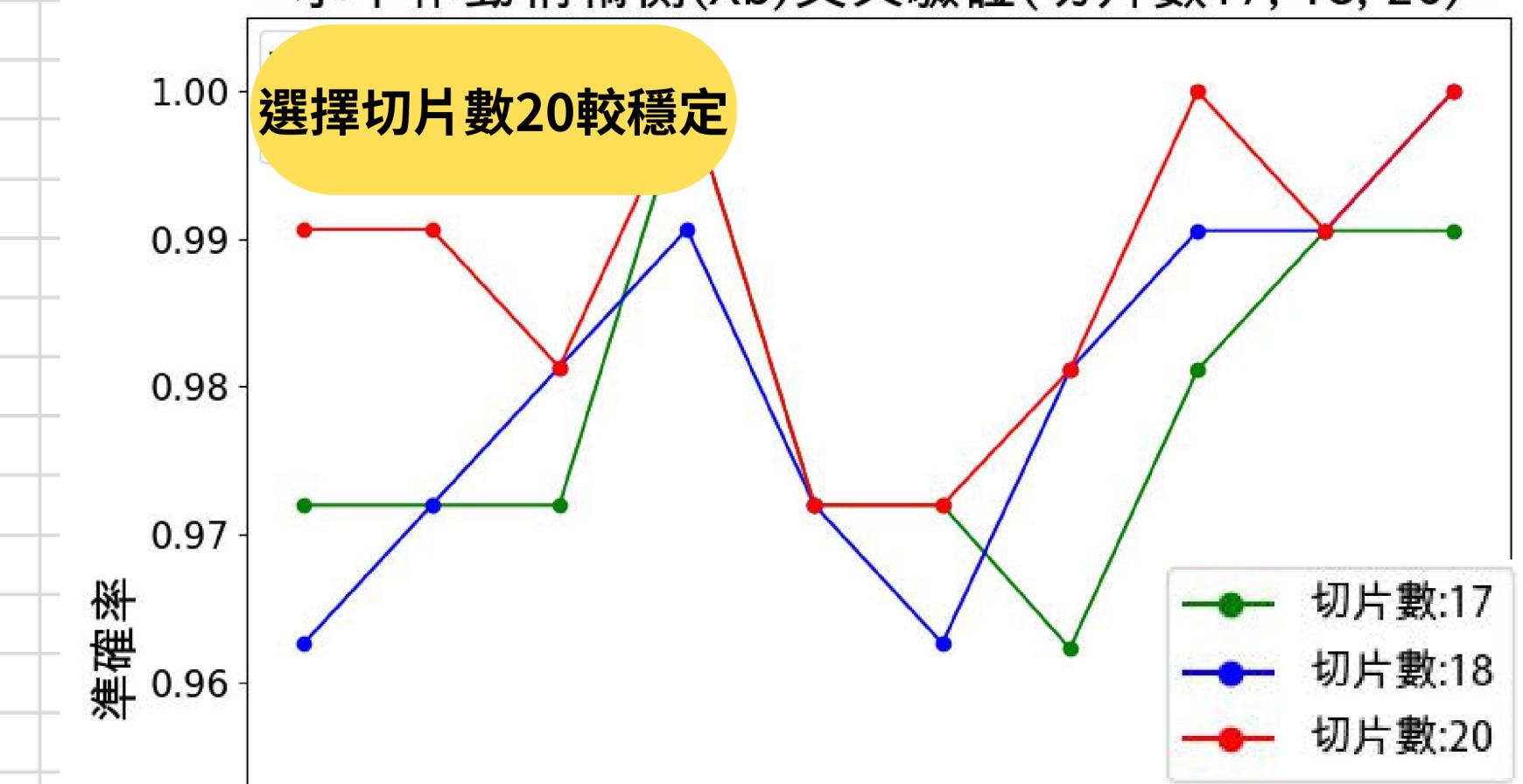
頻率中段特徵重要性高



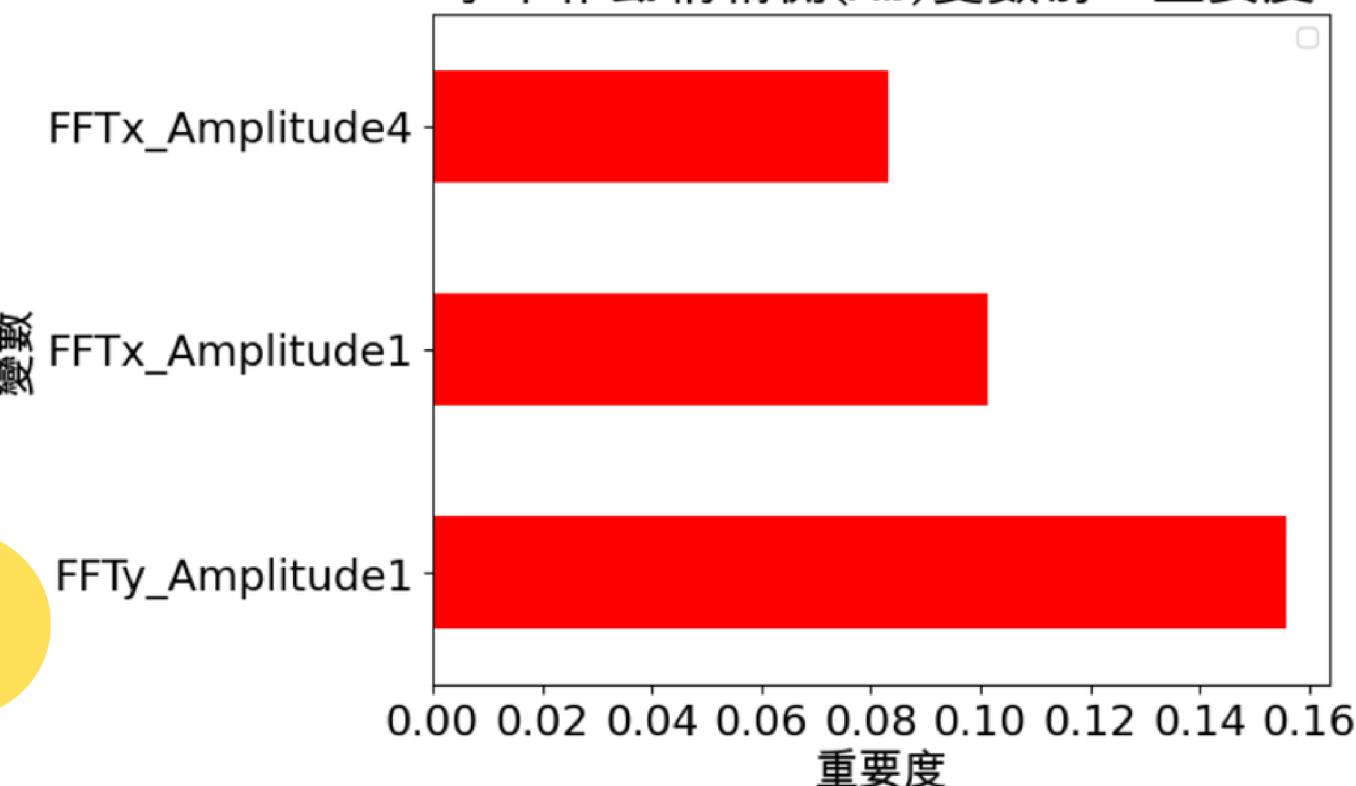
水平作動惰輪側(Xb)不同切片數下的準確率線圖



水平作動惰輪側(Xb)交叉驗證(切片數17, 18, 20)

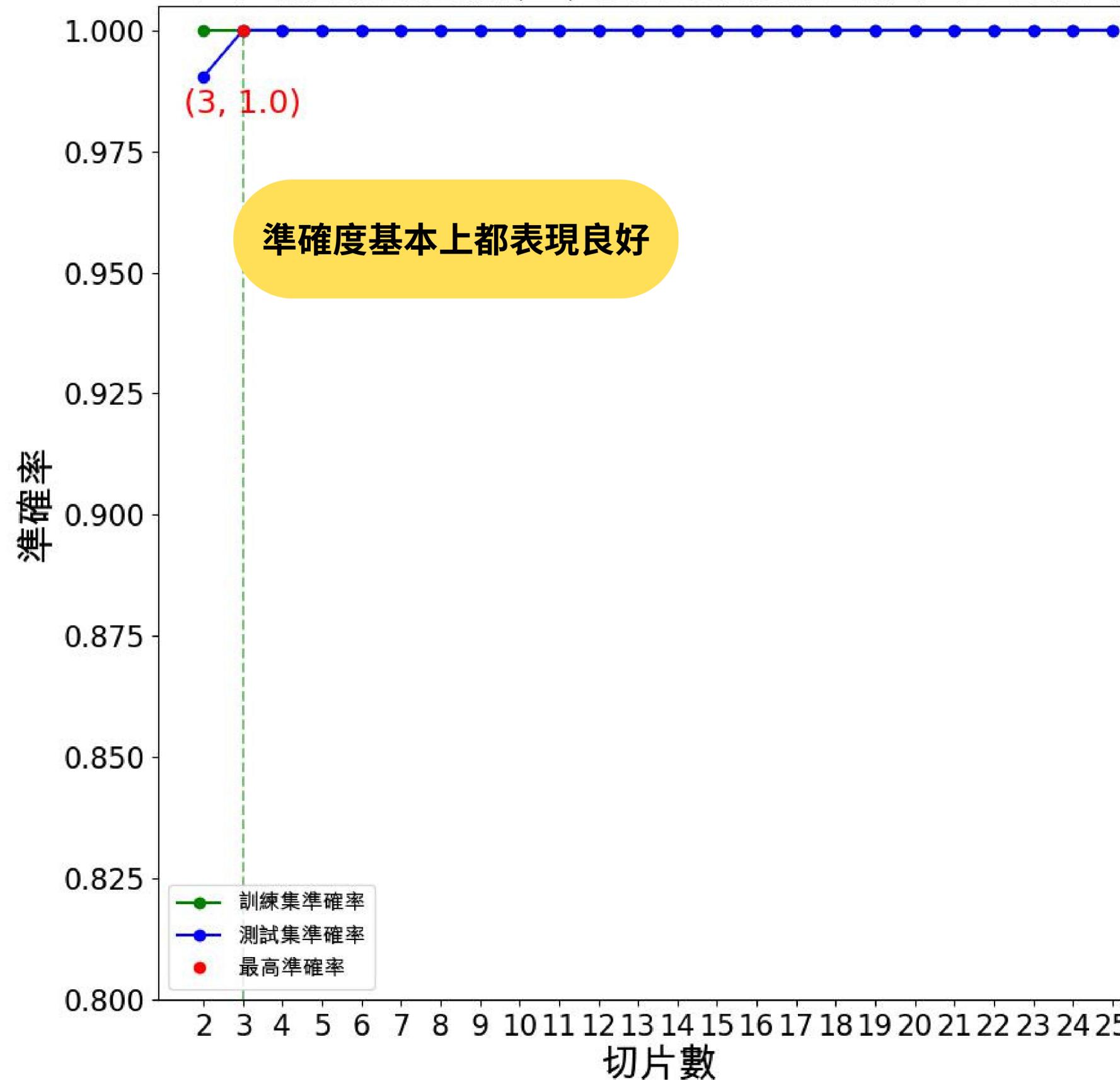


水平作動惰輪側(Xb)變數前三重要度



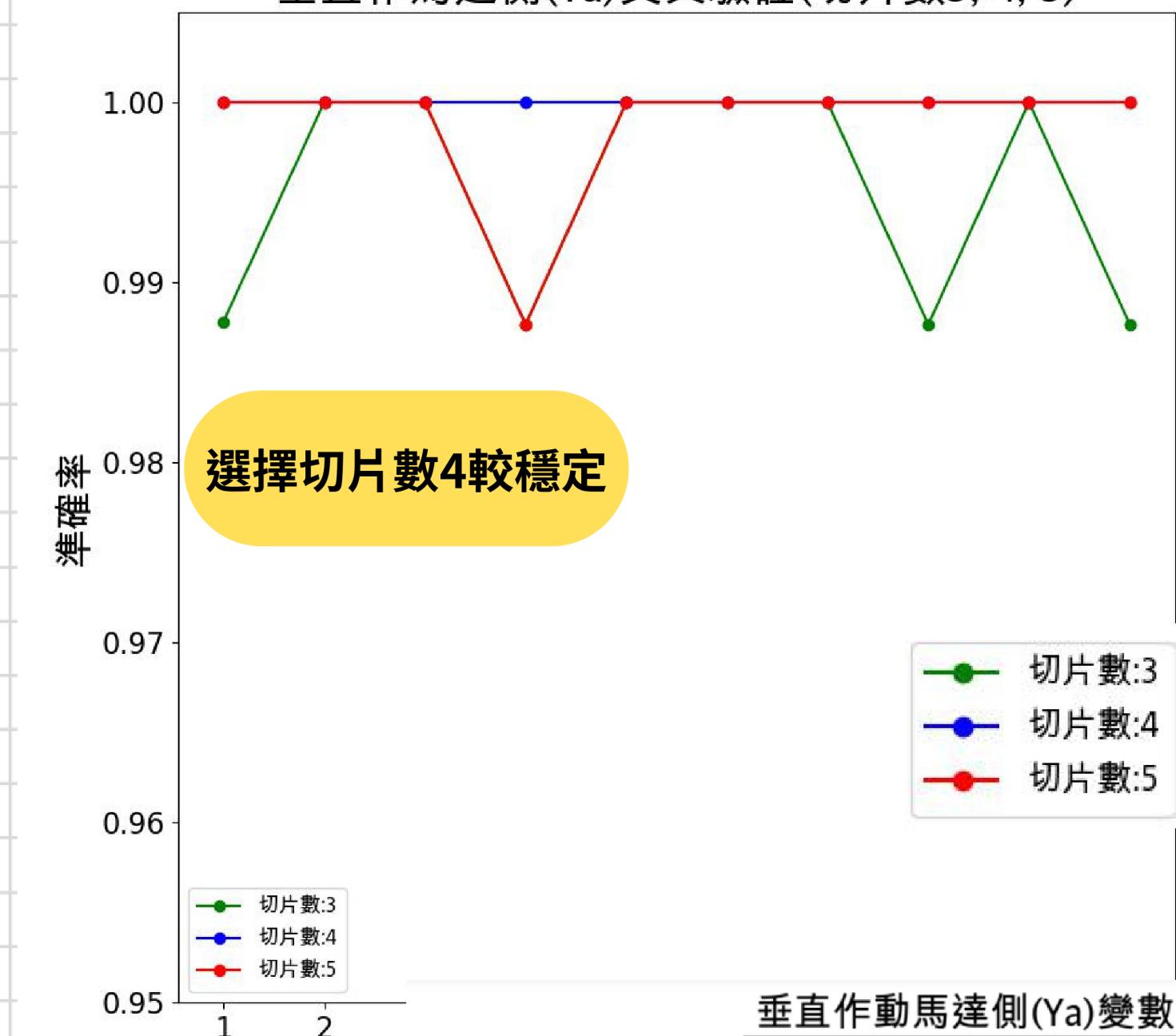
頻率前段特徵重要性高

垂直作動馬達側(Ya)不同切片數下的準確率線圖

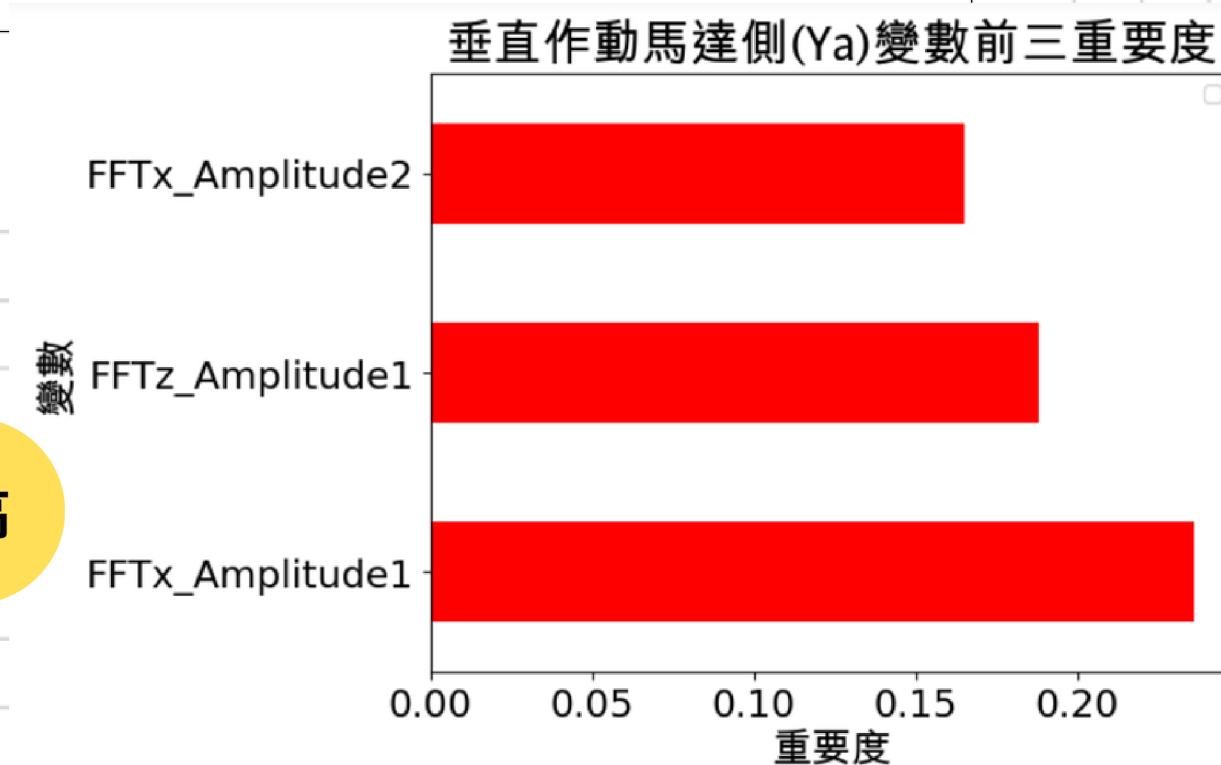


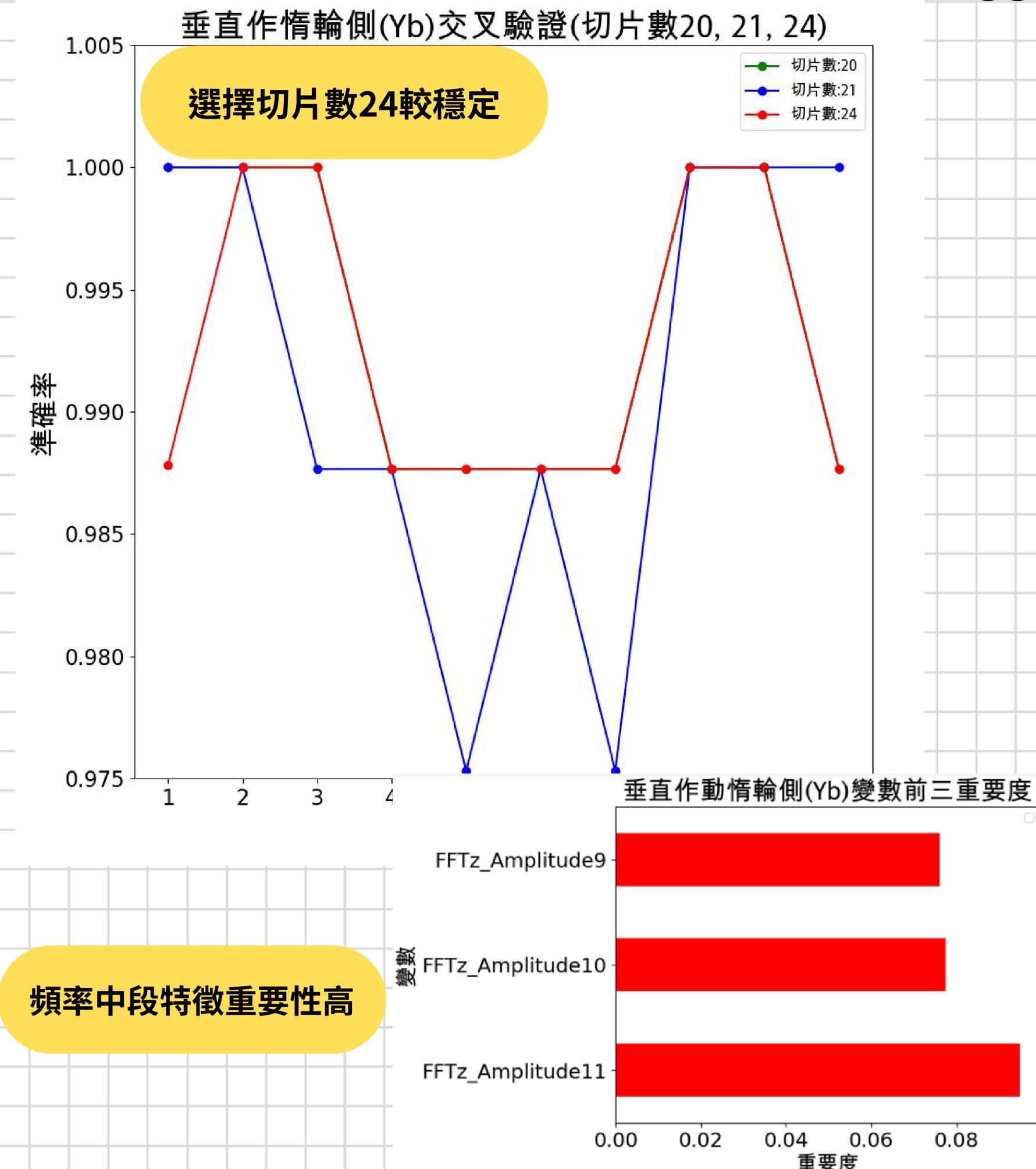
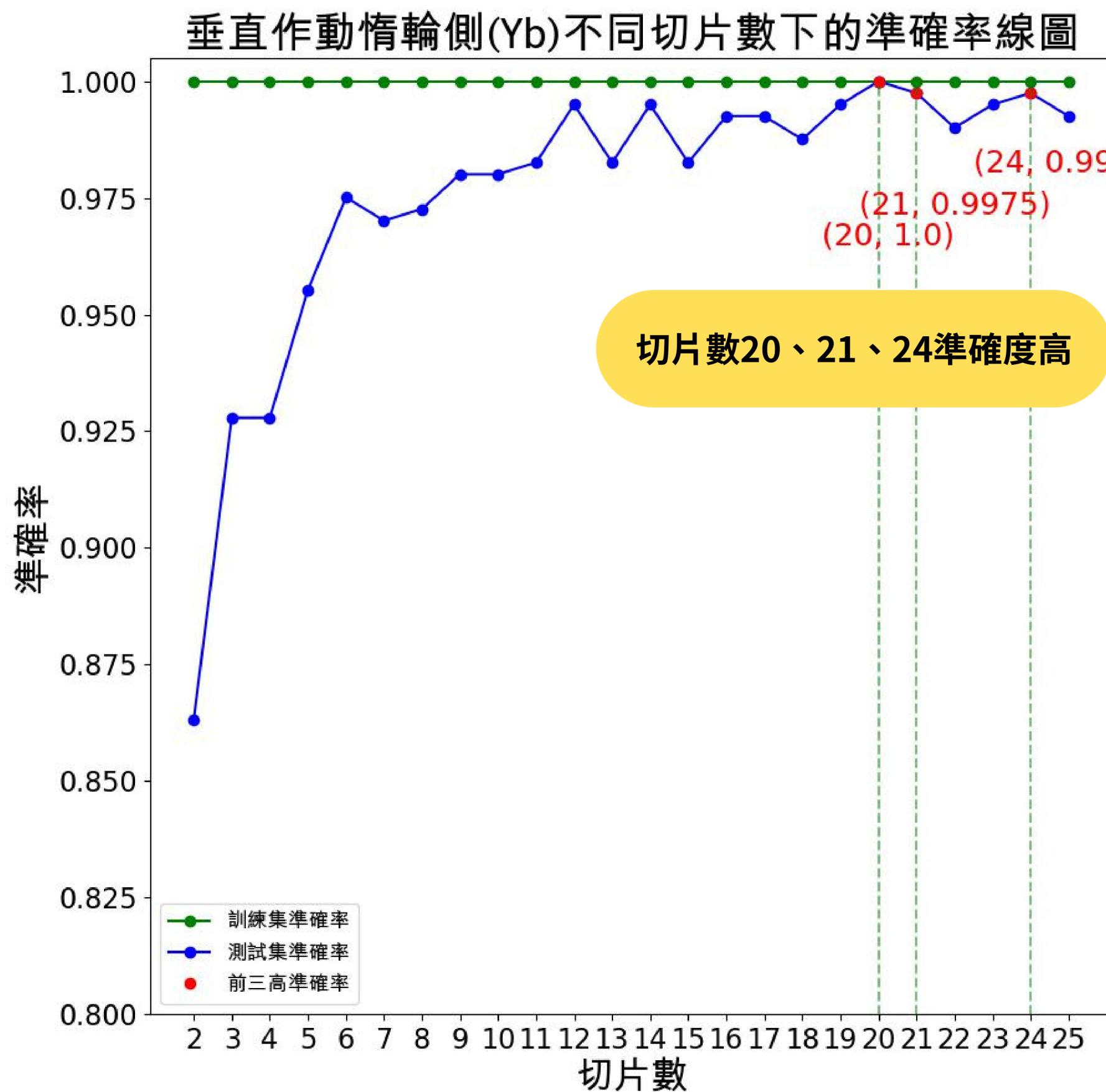
• 切片數為3、4、5的準確率較好

垂直作動馬達側(Ya)交叉驗證(切片數3, 4, 5)



頻率前段特徵重要性高

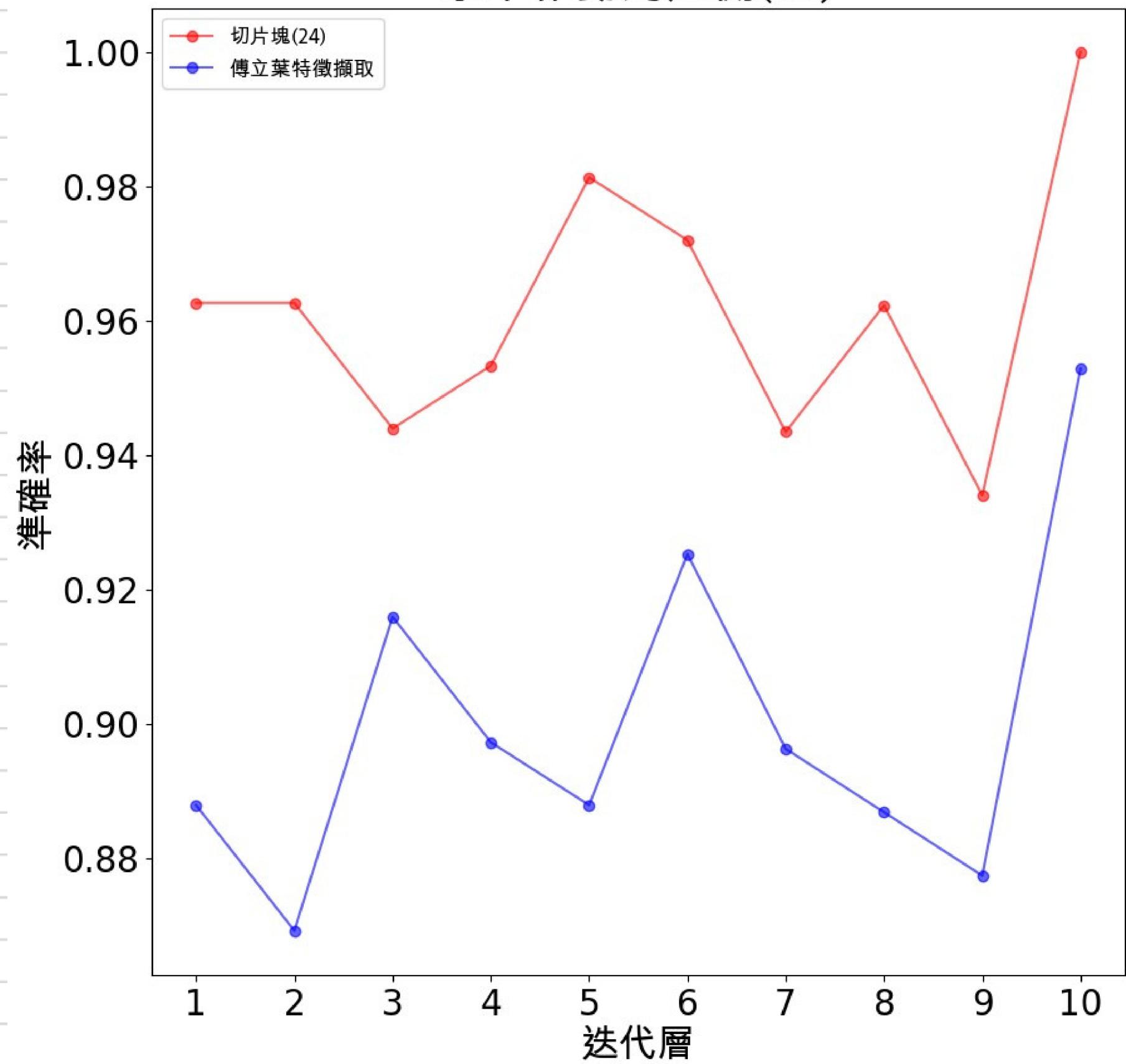




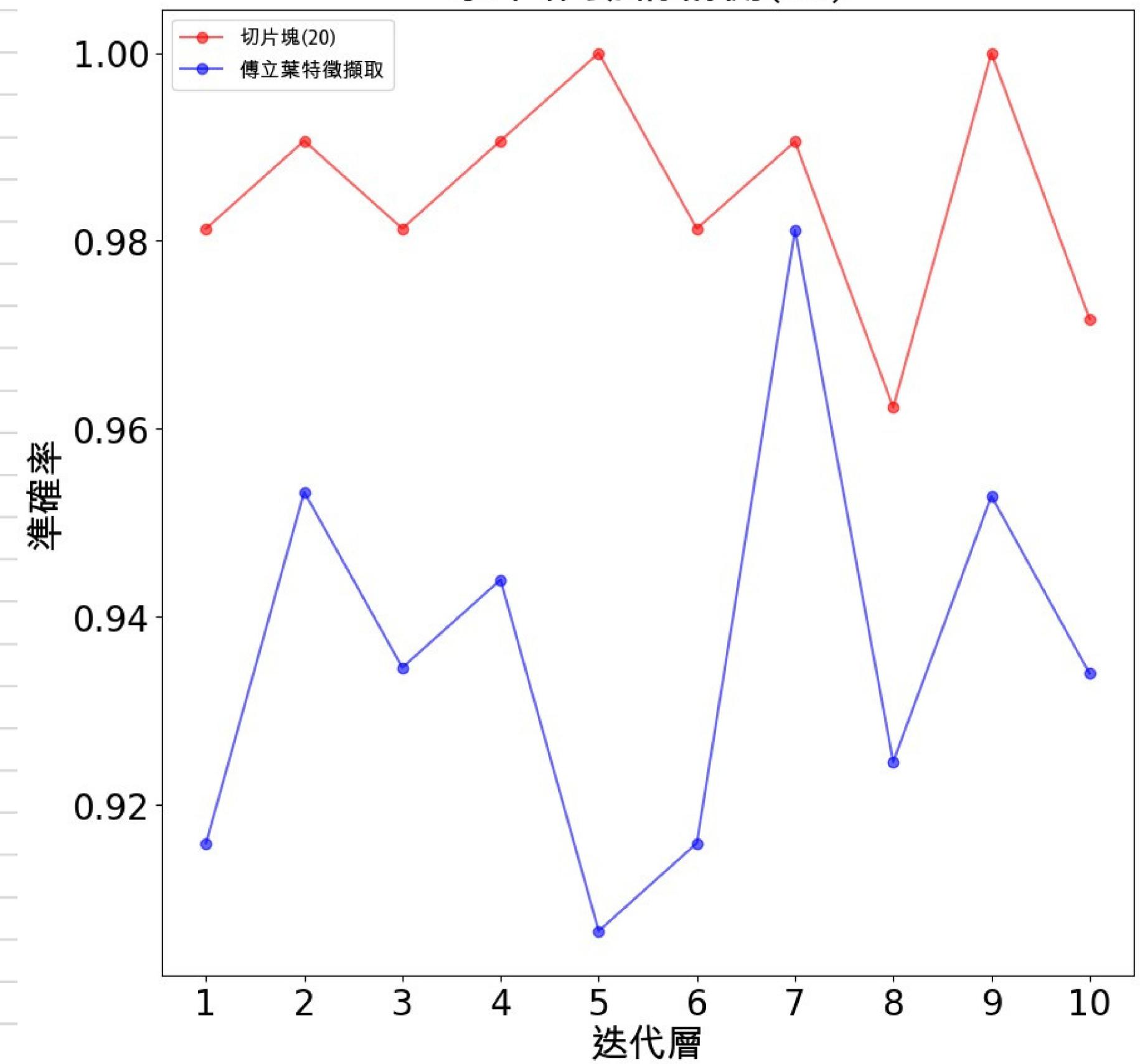


特徵擷取比較

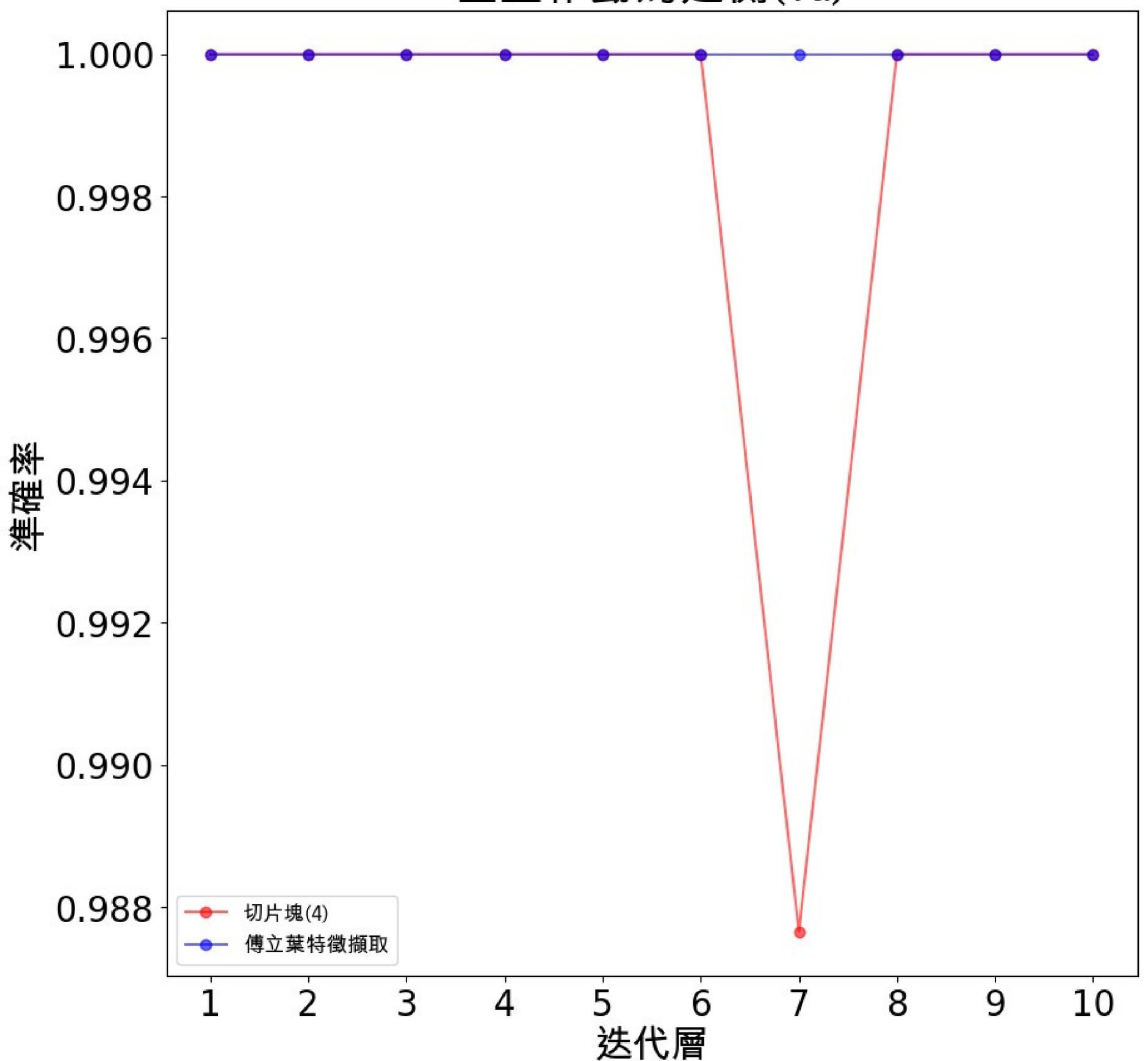
水平作動馬達側(Xa)



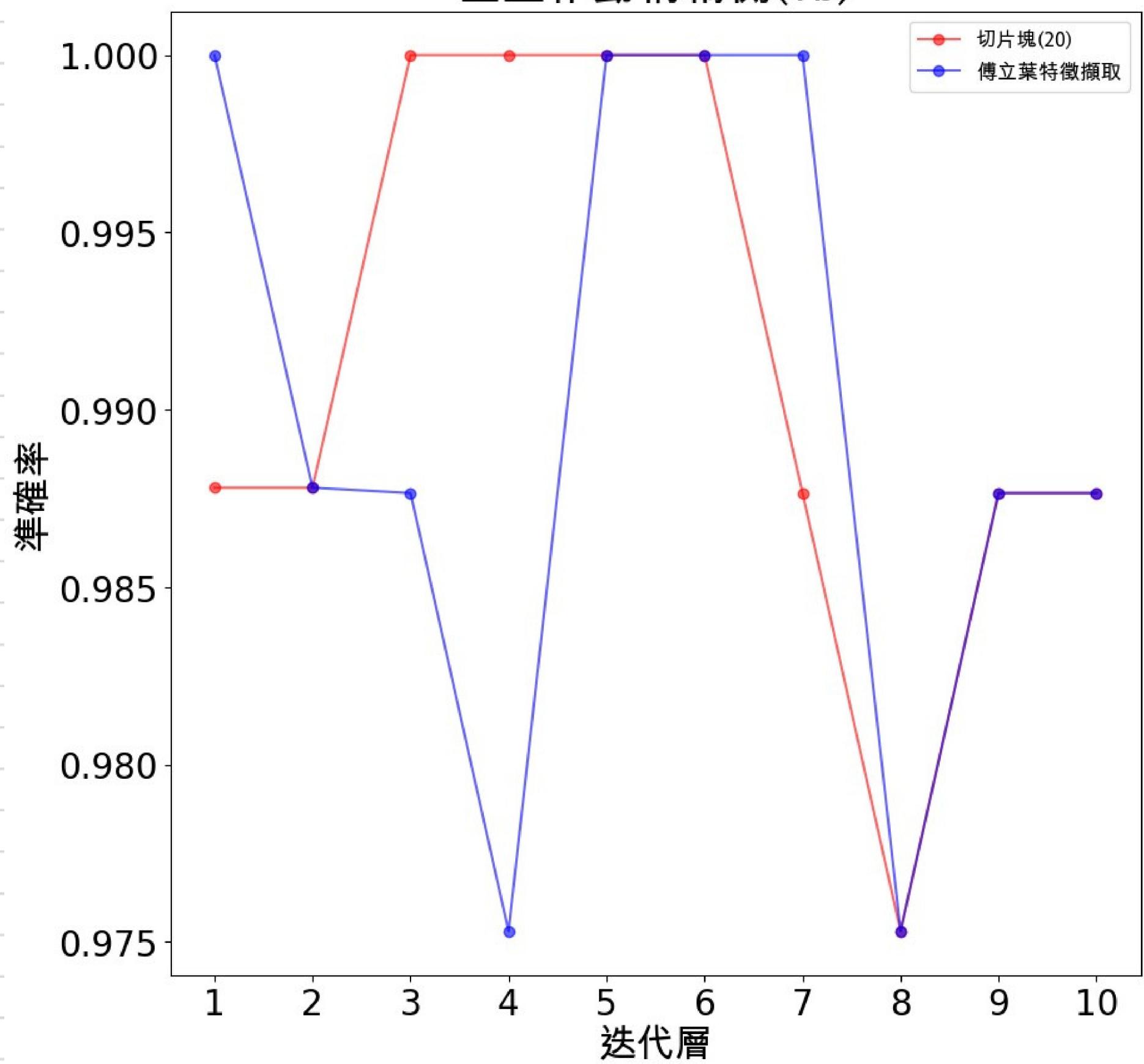
水平作動惰輪側(Xb)

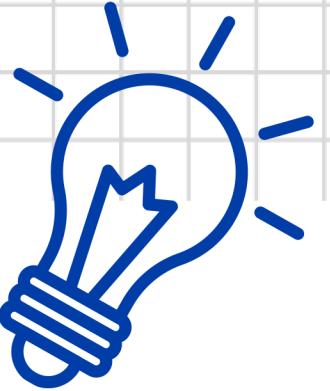


垂直作動馬達側(Ya)



垂直作動惰輪側(Yb)





建議決策方式

- 分段擷取極大值在各位置都有較穩定的效果，所以後續特徵擷取選擇此法
- 水平作動適合在惰輪側 (X_b) 做異常偵測
- 垂直作動適合在馬達側 (Y_a) 做異常偵測
- 分段擷取準確率較高原因：特徵較為細緻
- 未來展望：
 1. 根據波形特徵的變數重要度，可以在分段擷取時增加有效變數
 2. 分段是否為準確率上升的唯一原因？還是變數的選擇？

輔助決策工具設計稿

機械手臂振動情況與健康度

選擇作動位置:

水平作動下馬達側(Xa)

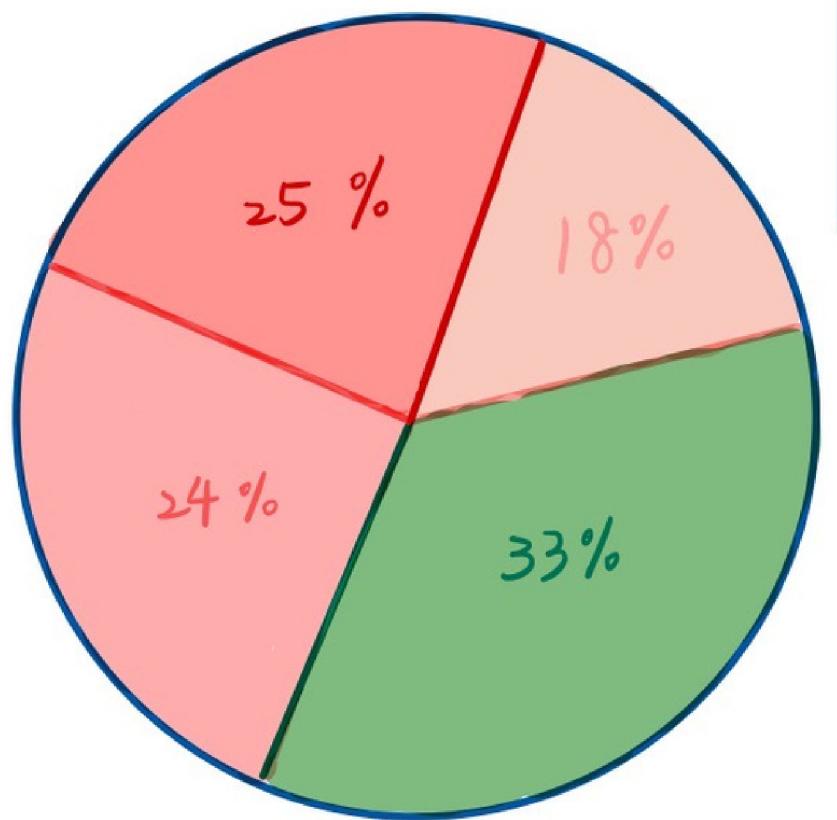
上傳文件(至多25個):

Browse... No file selected

處理文件

可選擇25個資料 (.txt)

水平作動下馬達側 (Xa) 負荷圓餅圖



健康度
64 %

正常	異常
16	9

水平作動下馬達側 (Xa) 負荷一覽表

■ 65	■ 80	■ 95	■ 130	■ 65
■ 65	■ 80	■ 95	■ 130	■ 65
■ 65	■ 80	■ 95	■ 130	■ 65
■ 65	■ 80	■ 95	■ 130	■ 65
■ 65	■ 80	■ 95	■ 130	■ 65

檔案名稱: gvb 0005

負荷: 80

檔案名稱: gvb 00015

負荷: 65

檔案名稱: gvb 00020

負荷: 95

檔案名稱: gvb 00025

負荷: 130



Q&A