

從聲音訊號到故障預測：利用機器學習實現設備預測性維護

製造數據科學 Group G

R13725051 陳怡均 R13725057 廖婕妤 B11705056 黃雋亞 B11703063 鄭凱元 B10704026 劉子揚





Table Of Contents

1

Motivation & Background

2

Data Collection

3

Methodology

4

Result Analysis & Discussion

Motivation & Background

痛點分析

故障判斷主觀

傳統設備故障診斷依賴人工經驗，容易受到操作人員的主觀判斷影響，可能導致診斷結果不一致或錯誤，降低維修效率與準確性

成本高昂

人工檢查設備需要專業知識與大量時間，再加上延遲的故障檢測可能導致生產線停機，產生高昂的停工成本

異常無法及時發現

傳統監控手段不見得能夠辨識所有潛在異常，如設備發出聲音的細微變化，常常是機器已經出現問題後才發現異狀

無法有效累積資料

若沒有系統性收集資料，將無法建立訓練的基準，也不能追蹤長期趨勢與機台健康診斷

→ 預測性維護能降低成本、提高生產效率

Motivation & Background

背景

反應性維護

即時性低，可能造成無法預期的停機損失

預防性維護

可能導致過度維護，增加不必要的成本

預測性維護

透過感測器監測設備運行數據，並結合資料分析技術，預測設備潛在故障，進而在故障發生前進行維修，有效減少停機時間並最佳化資源配置

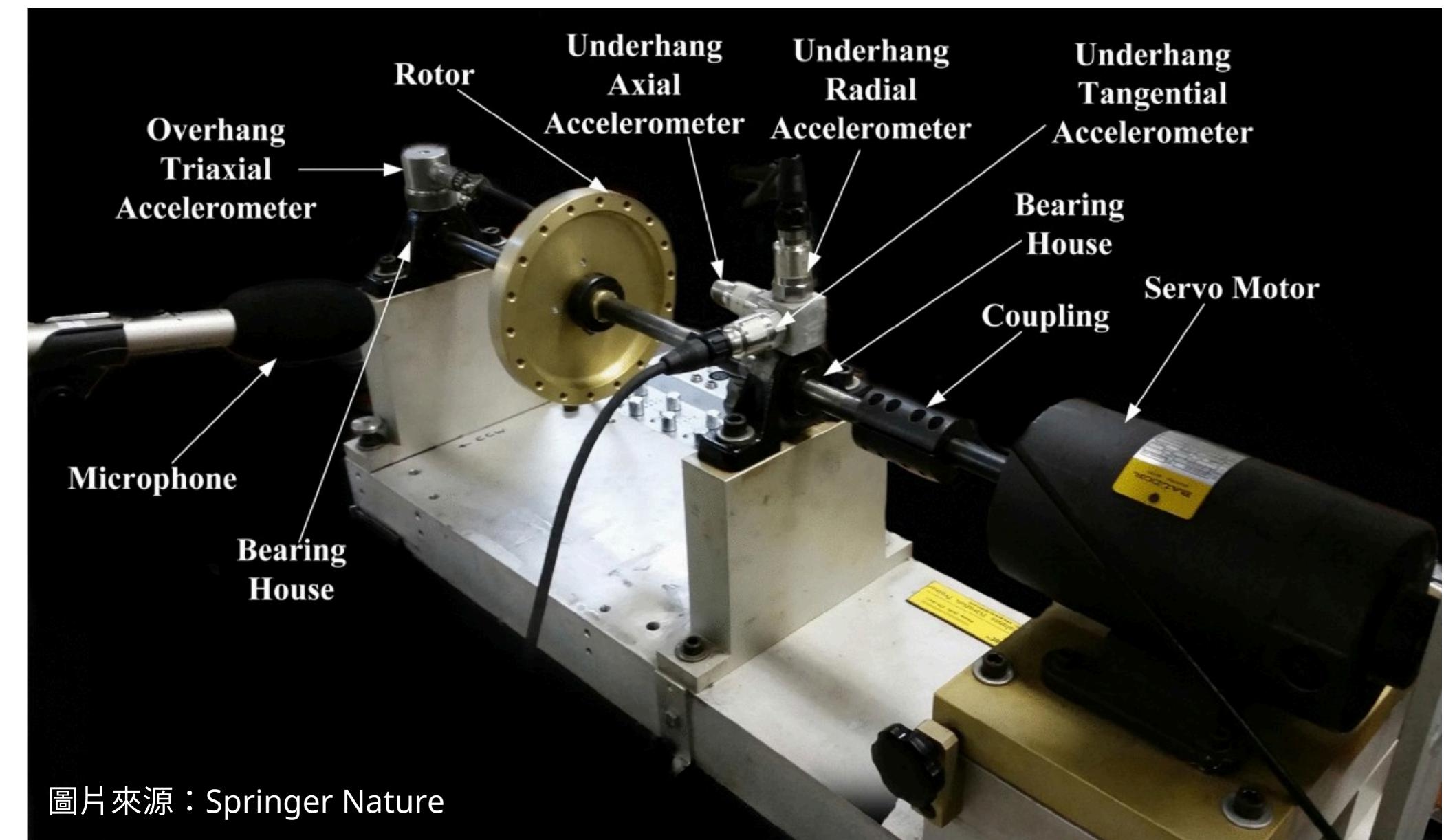
問題定義

運用深度學習 CNN 與傳統機器學習 Random Forest 建立機械設備所發出音訊之故障偵測模型

Data | MAFAULDA Machinery Fault Database

資料概述

透過機械故障模擬器取得之 1951 筆多變量時間序列資料，共有六種故障類型，部分類型包含數種故障元件。每筆資料皆以 50kHz 取樣五秒，共 250,000 列。



Data | MAFAULDA Machinery Fault Database

資料欄位

共有8個欄位，分別是轉速表訊號、下懸軸承的軸向、徑向、切向的加速計、上懸軸承的軸向、徑向、切向的加速計與麥克風訊號。

標籤與資料量

標籤	資料筆數		
正常	49		
水平未對準	197	上懸軸承	
垂直未對準	301	上懸軸承保持架故障	188
不平衡	333	上懸軸承外圈故障	188
		上懸軸承滾珠故障	137
下懸軸承			
下懸軸承保持架故障	188		
下懸軸承外圈故障	184		
下懸軸承滾珠故障	186		
		總計	1951

模型建立流程



方法一 | CNN

資料前處理

• 標籤建立

Label Index	類別名稱	說明
0	normal	正常狀況聲音
1	imbalance	軸心不平衡
2	horizontal-misalignment	水平未對準
3	vertical-misalignment	垂直未對準

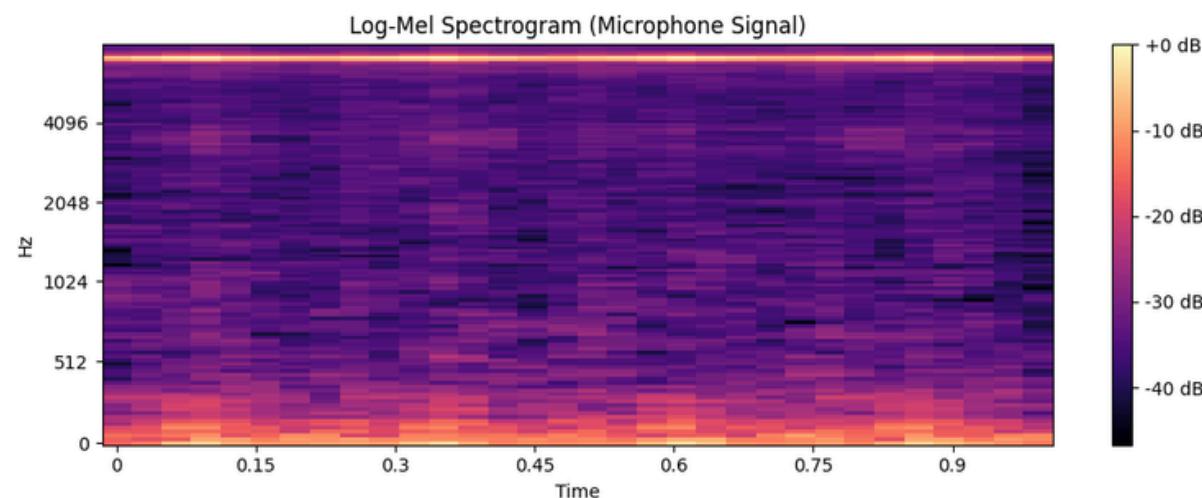
Label Index	類別名稱	說明
4	overhang_fault1	上懸軸承保持架故障
5	overhang_fault2	上懸軸承外圈故障
6	overhang_fault3	上懸軸承滾珠故障
7	underhang_fault1	下懸軸承保持架故障
8	underhang_fault2	下懸軸承外圈故障
9	underhang_fault3	下懸軸承滾珠故障



方法一 | CNN

資料前處理

- 音訊切片 (Segmentation)：將長音訊切成固定長度片段供模型輸入
- 資料增強 (Data Augmentation)：加入隨機白噪音、隨機增減音量
- 特徵提取 (Feature Extraction)：
 - 轉換 1D 音訊為 2D 特徵圖，便於 CNN 處理
 - 輸出 shape 為 $(128, 256, 1)$
- 加入動態特徵 (Delta)：捕捉聲音的變化趨勢（速度 / 加速度）。
- 特徵標準化 (Z-score Normalization)：去除樣本間的能量差異

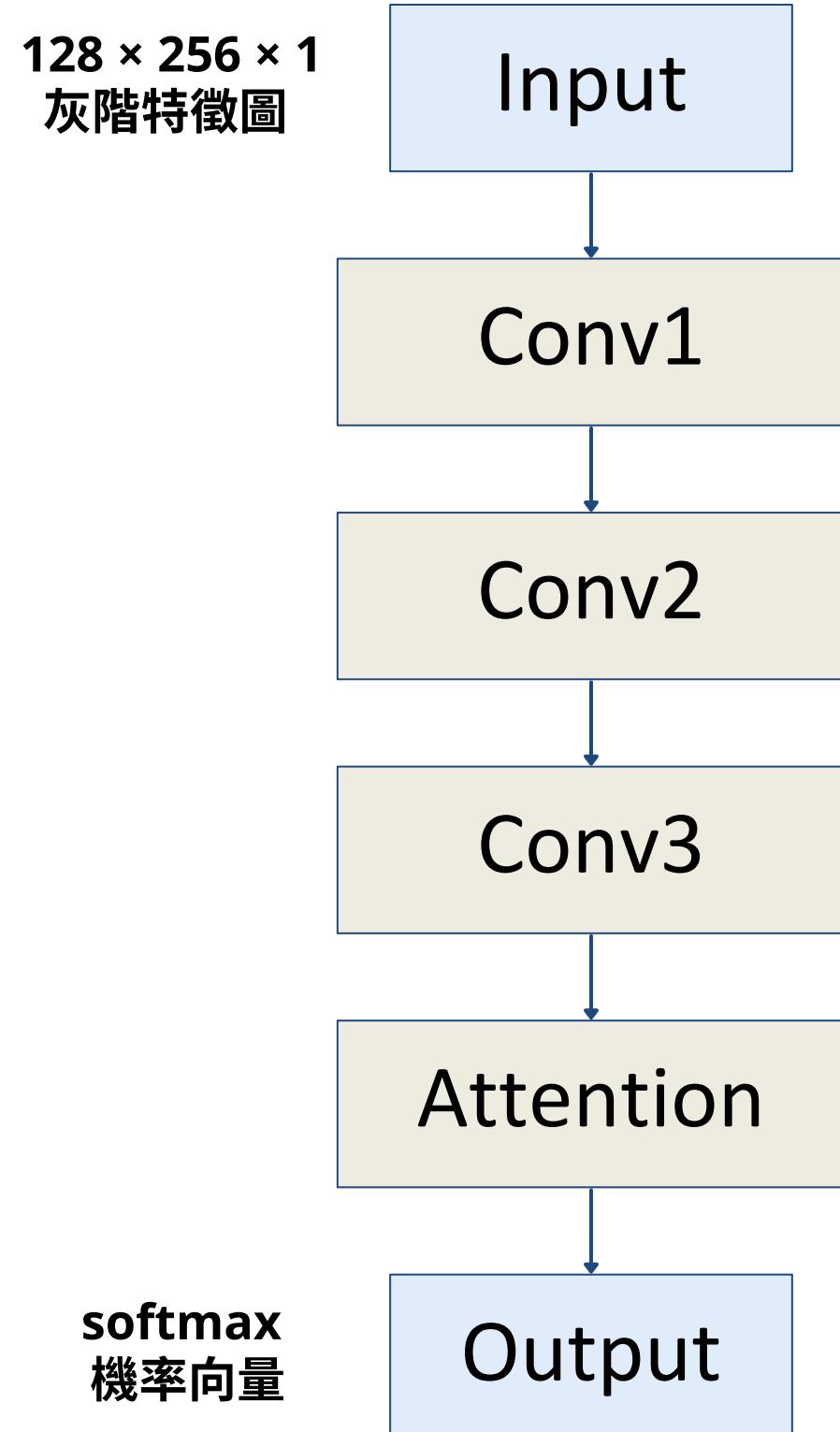


方法一 | CNN

模型架構

- CNN 提取局部特徵： $\text{Conv2D} \rightarrow \text{BN} \rightarrow \text{ReLU} \rightarrow \text{MaxPool}$ 3 次
- Reshape 成時間序列
- Attention 聚焦關鍵片段
 - 用 Dense + Tanh + Softmax 得到權重
- Dense + Dropout 轉為分類向量
- Softmax 輸出類別機率

‹#›

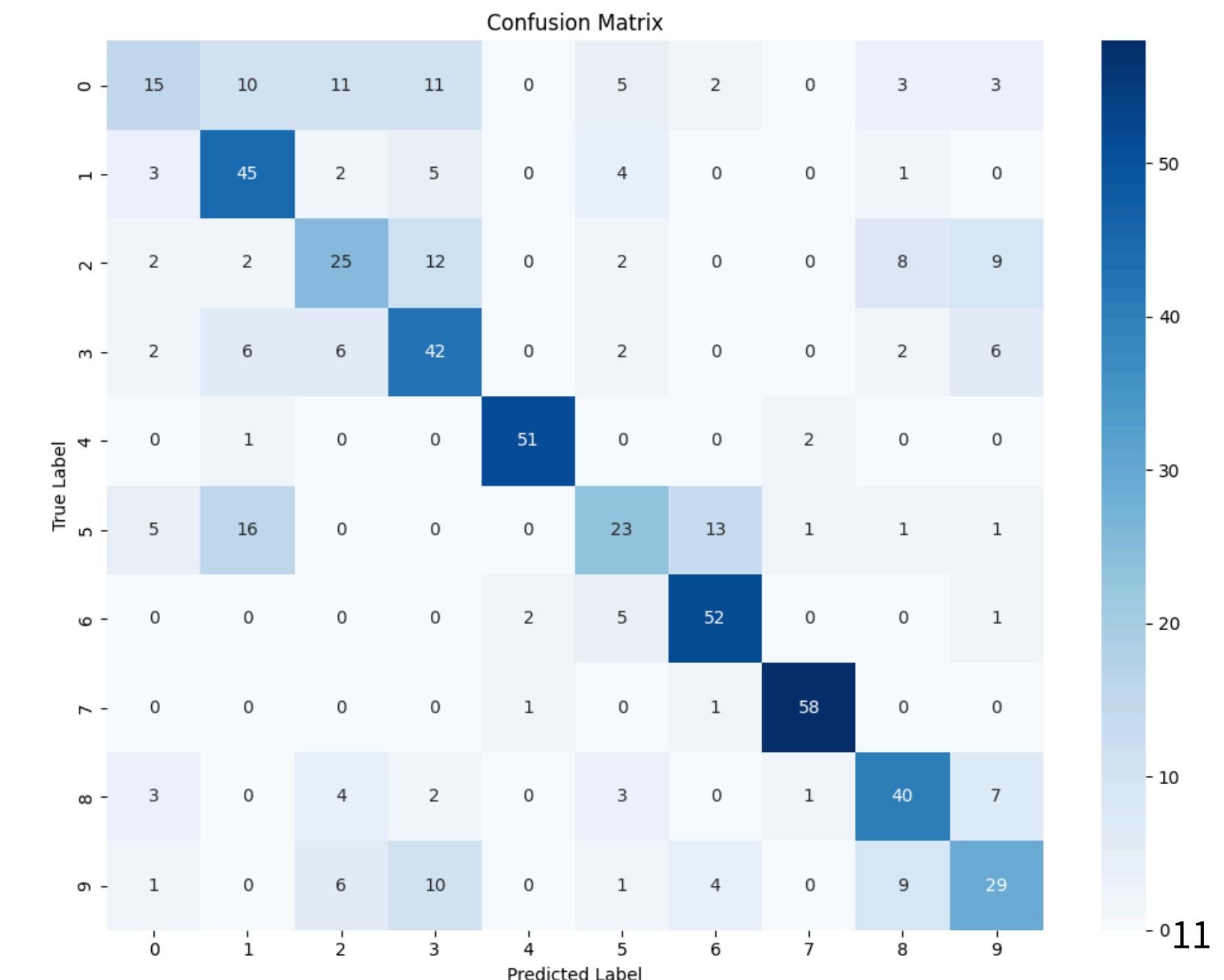


方法一 | CNN

模型表現

Classification Report:

	precision	recall	f1-score	support
0	0.4839	0.2500	0.3297	60
1	0.5625	0.7500	0.6429	60
2	0.4630	0.4167	0.4386	60
3	0.5122	0.6364	0.5676	66
4	0.9444	0.9444	0.9444	54
5	0.5111	0.3833	0.4381	60
6	0.7222	0.8667	0.7879	60
7	0.9355	0.9667	0.9508	60
8	0.6250	0.6667	0.6452	60
9	0.5179	0.4833	0.5000	60
accuracy			0.6333	600



方法二 | CNN + BiLSTM

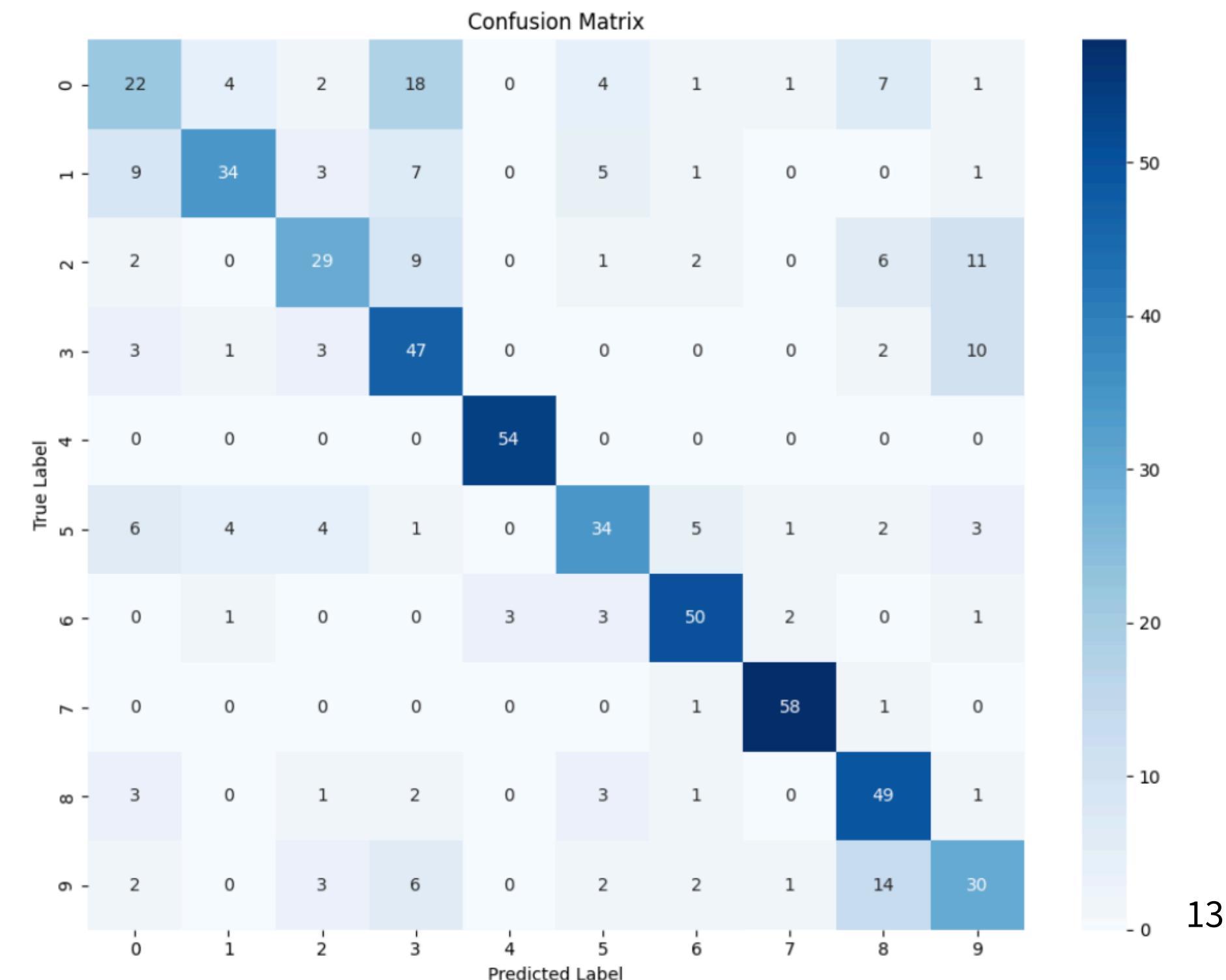
模型架構



方法二 | CNN + BiLSTM

模型表現

	precision	recall	f1-score	support
0	0.4681	0.3667	0.4112	60
1	0.7727	0.5667	0.6538	60
2	0.6444	0.4833	0.5524	60
3	0.5222	0.7121	0.6026	66
4	0.9474	1.0000	0.9730	54
5	0.6538	0.5667	0.6071	60
6	0.7937	0.8333	0.8130	60
7	0.9206	0.9667	0.9431	60
8	0.6049	0.8167	0.6950	60
9	0.5172	0.5000	0.5085	60
accuracy		0.6783		600



小結 | CNN / CNN + BiLSTM

分類結果探討

嘗試改善資料前處理方法與模型架構，但分類準確率僅維持在 60% 左右，無法準確分類所有類別。

檢視原始資料：

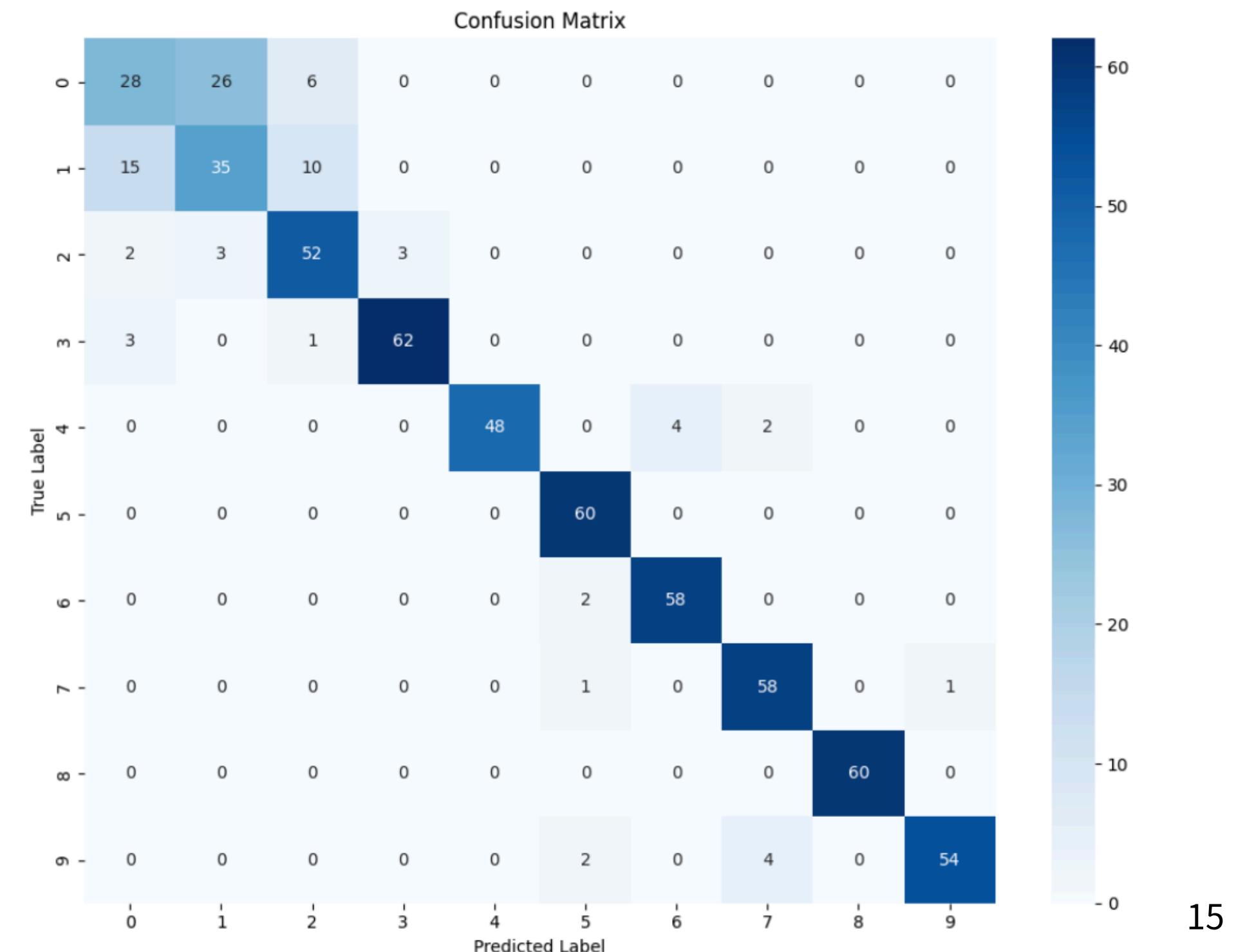
→ 若讓模型只針對第八欄位麥克風音訊學習，能更準確地學習資料集的特徵

小結 | CNN / CNN + BiLSTM

模型表現

	precision	recall	f1-score	support
0	0.5833	0.4667	0.5185	60
1	0.5469	0.5833	0.5645	60
2	0.7536	0.8667	0.8062	60
3	0.9538	0.9394	0.9466	66
4	1.0000	0.8889	0.9412	54
5	0.9231	1.0000	0.9600	60
6	0.9355	0.9667	0.9508	60
7	0.9062	0.9667	0.9355	60
8	1.0000	1.0000	1.0000	60
9	0.9818	0.9000	0.9391	60

accuracy



方法三 | Random Forest 1

特徵工程

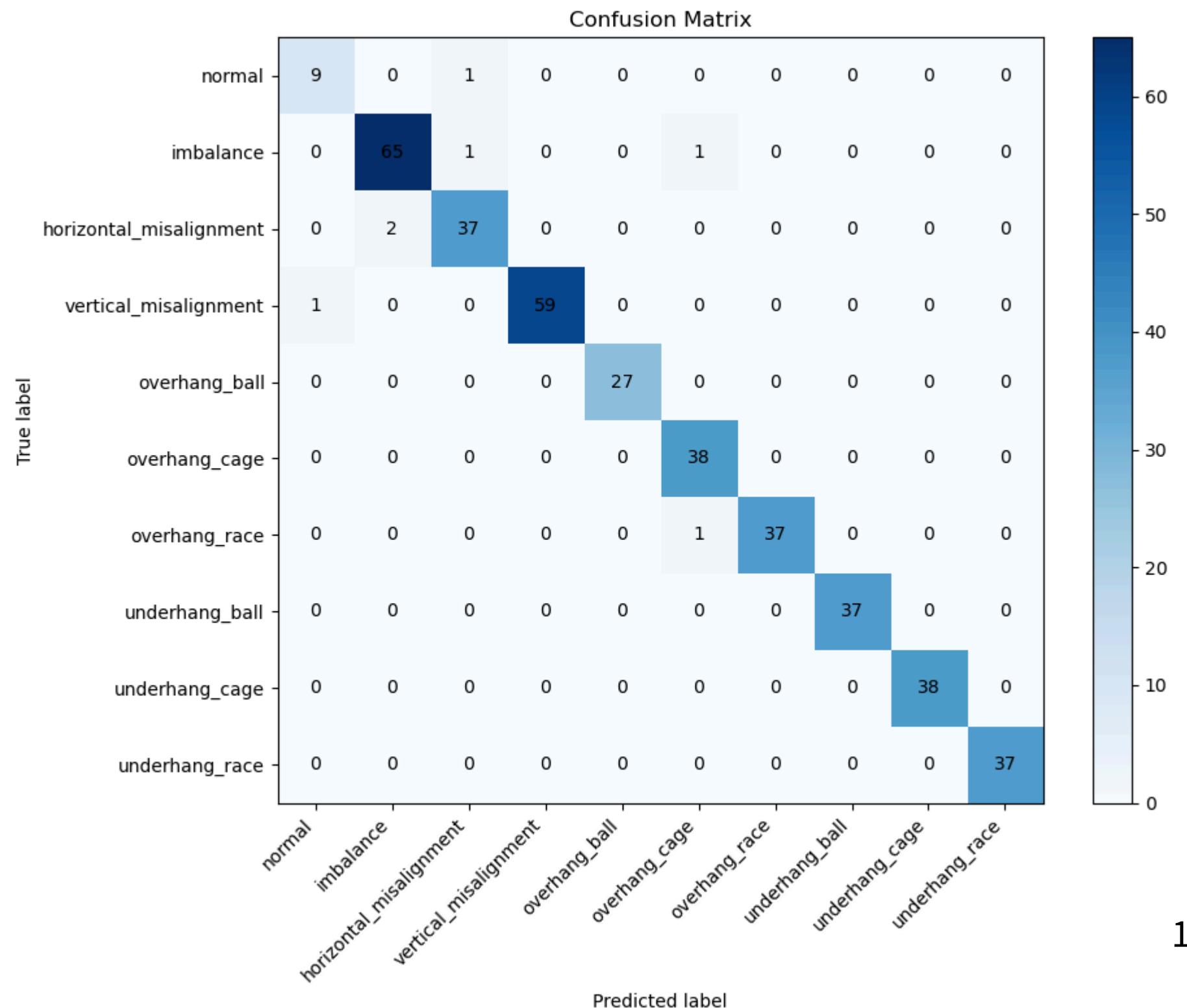
1. 平均
2. 標準差
3. 最小值
4. 最大值
5. 中位數

資料切割

Train / Test : 80 / 20

模型表現

Accuracy: 0.98



方法三 | Random Forest 1

Cross Validation

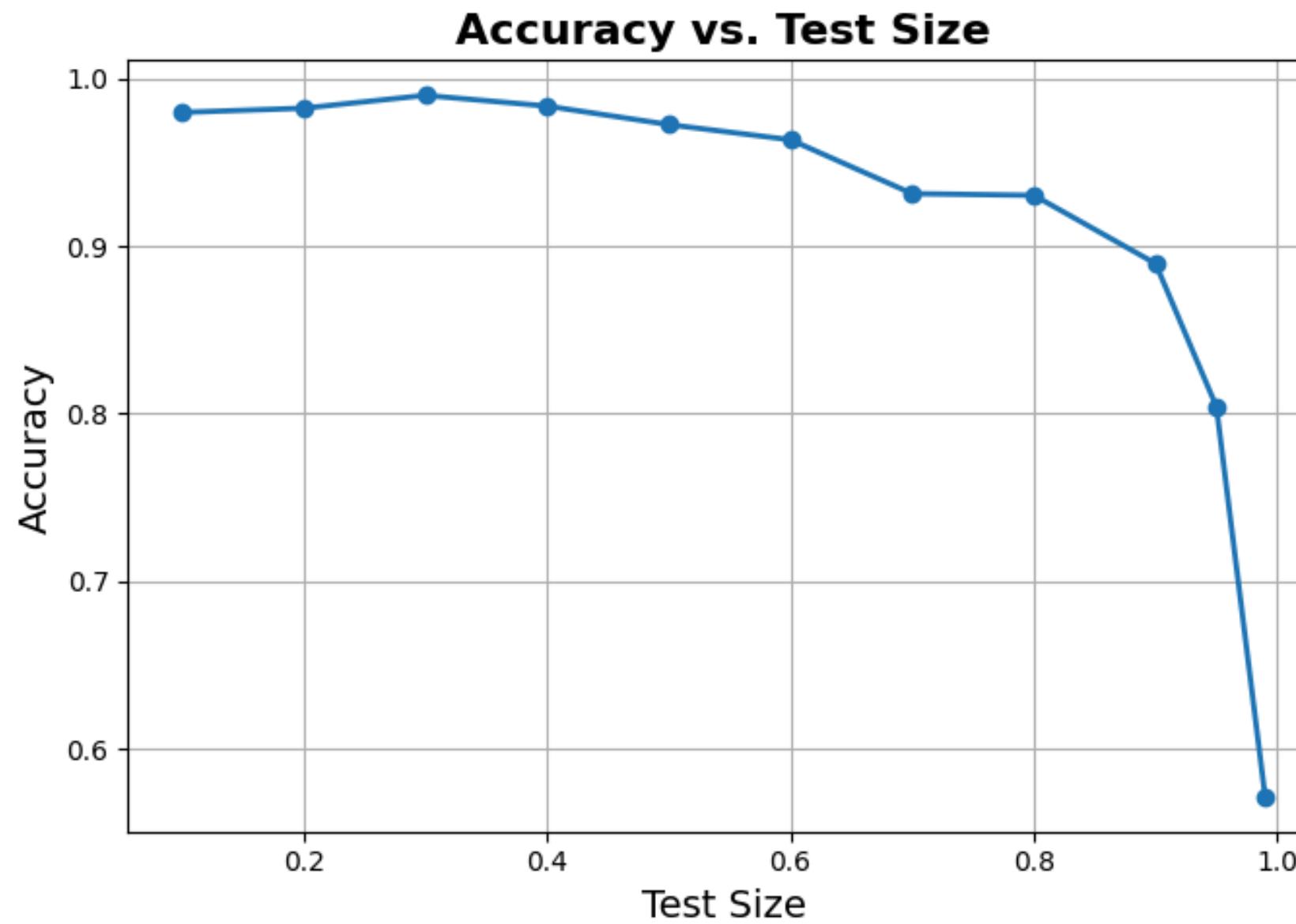
Accuracy	
Fold	
0	0.987212
1	0.992308
2	0.982051
3	0.987179
4	0.987179

Fault Type Prediction Result

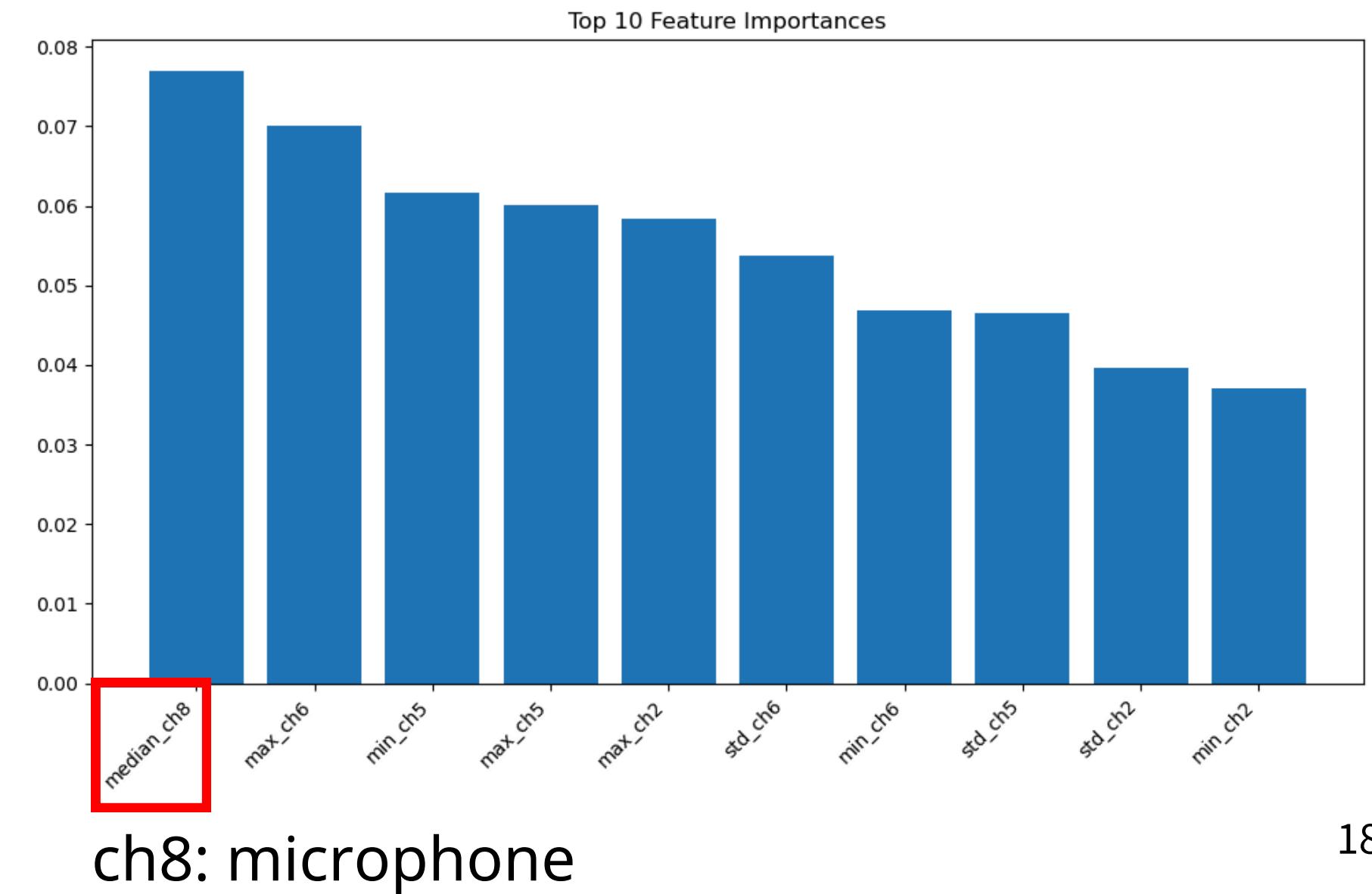
		precision	recall	f1-score	support
	normal	0.90	0.90	0.90	10
	imbalance	0.97	0.97	0.97	67
	horizontal_misalignment	0.95	0.95	0.95	39
	vertical_misalignment	1.00	0.98	0.99	60
	overhang_ball	1.00	1.00	1.00	27
	overhang_cage	0.95	1.00	0.97	38
	overhang_race	1.00	0.97	0.99	38
	underhang_ball	1.00	1.00	1.00	37
	underhang_cage	1.00	1.00	1.00	38
	underhang_race	1.00	1.00	1.00	37
	accuracy			0.98	391
	macro avg	0.98	0.98	0.98	391
	weighted avg	0.98	0.98	0.98	391

方法三 | Random Forest 1

正常 Test Size 表現穩健



音訊確實是重要特徵



方法四 | Random Forest 2

特徵工程

頻率特徵 21 個 (轉速表以外的 7 個感測器 × 3 個諧波)

短時距傅立葉轉換 (取旋轉頻率1倍、2倍及3倍之頻譜幅值平均*)

統計特徵 64 個 (8 個感測器 × 8 個統計量)

平均值、標準差、最大值、最小值、均方根、偏度、峰度、峰值因數

資料切割

Train / Test : 80 / 20，另有採用交叉驗證

資料平衡

SMOTE

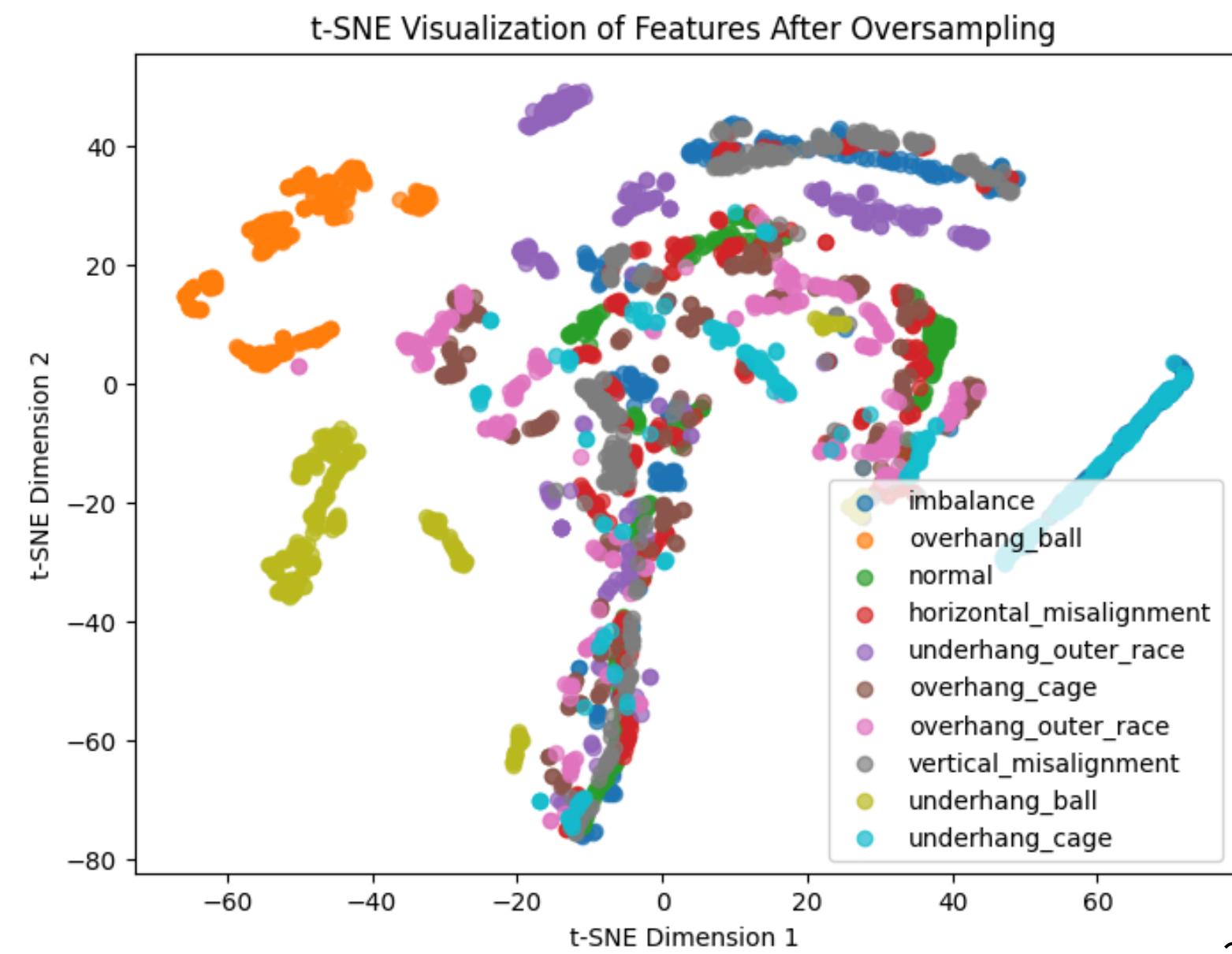
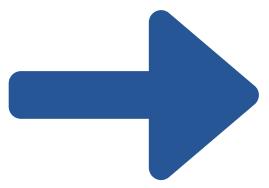
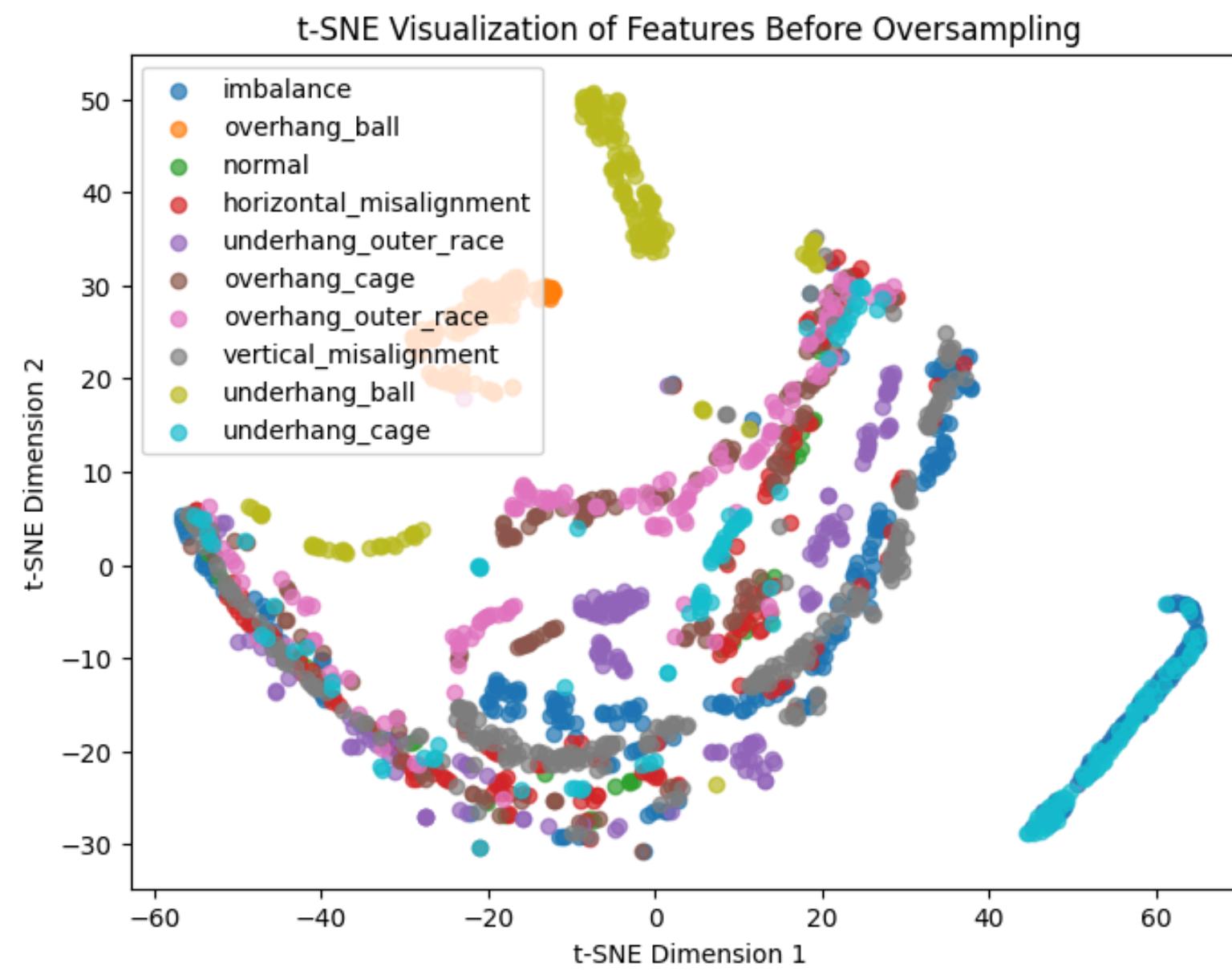
維度縮減

取 Permutation Importance 為正值的 22 個特徵

*參考 Marins, Matheus & Ribeiro, Felipe & Netto, Sergio & da Silva, Eduardo. (2017). Improved Similarity-Based Modeling for the Classification of Rotating-Machine Failures.

方法四 | Random Forest 2

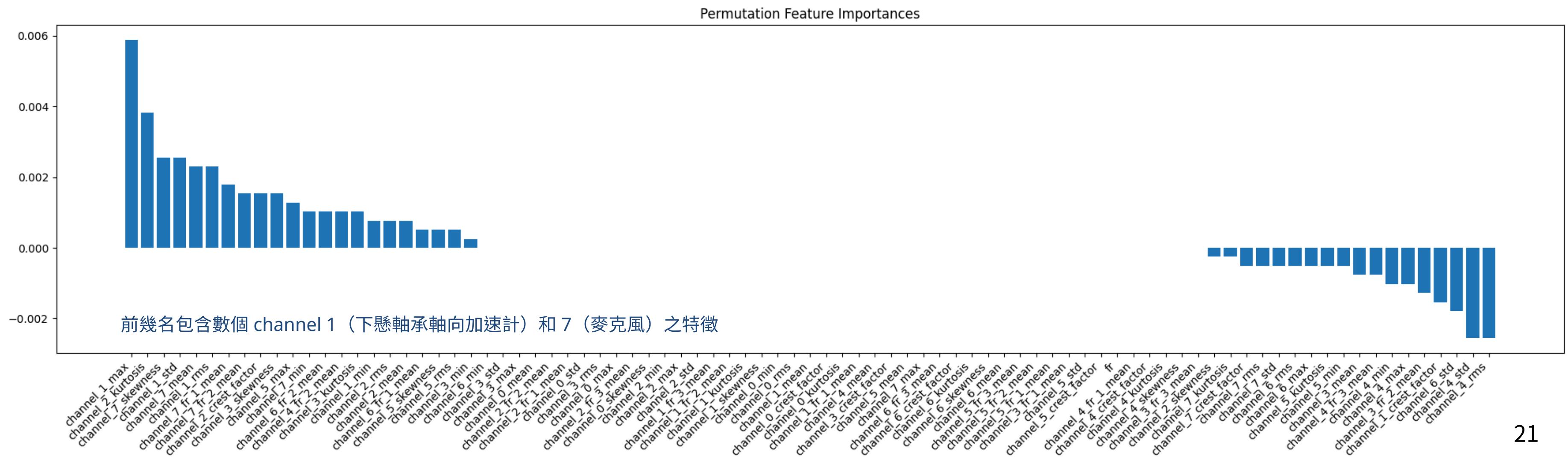
SMOTE 資料平衡前後之資料分布



方法四 | Random Forest 2

重要特徵篩選

篩選 Permutation Importance 值為正之特徵共 22 個，並重新訓練模型



方法四 | Random Forest 2

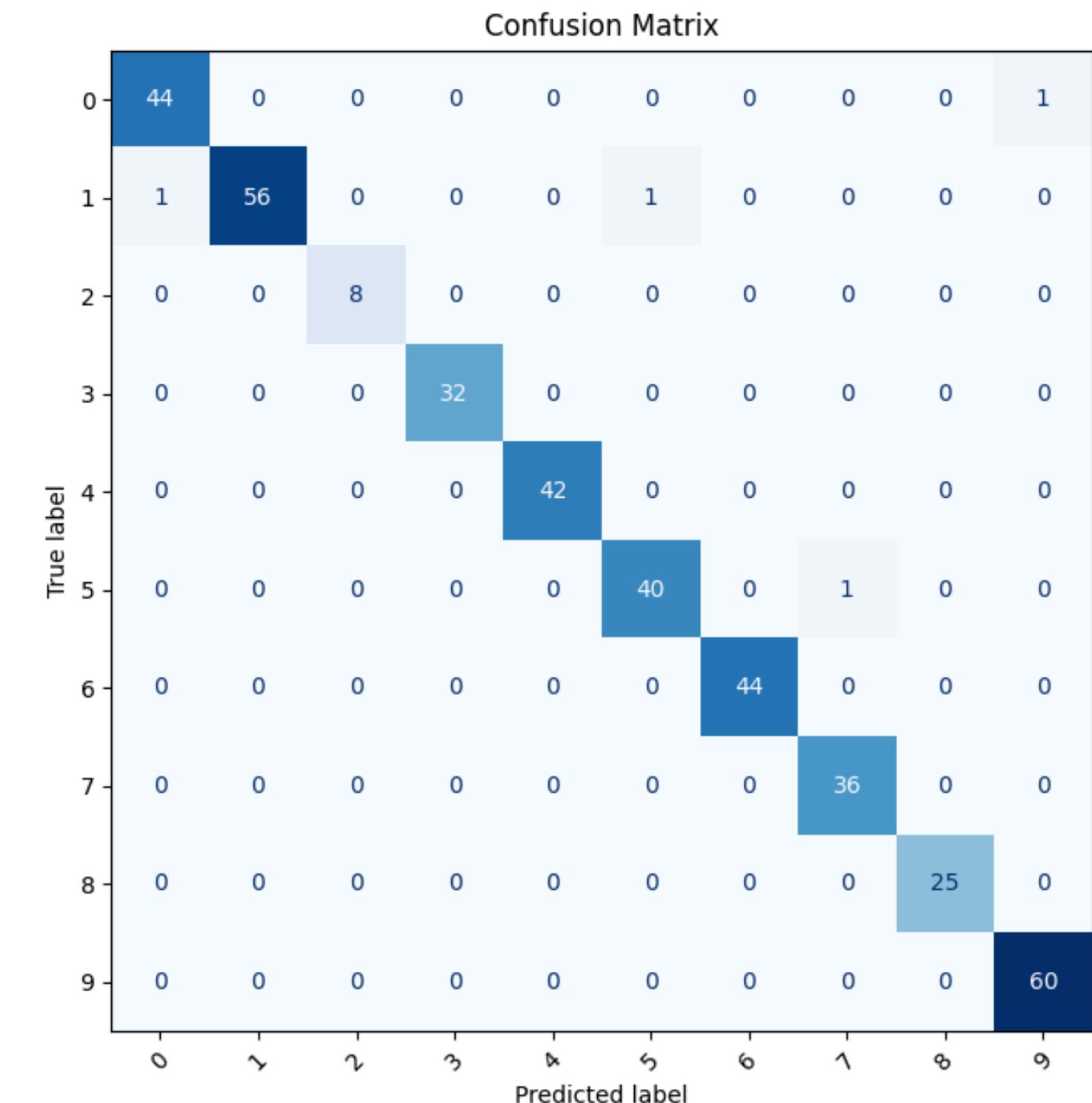
模型表現

Precision、Recall 及 F1-score
均達 99%

交叉驗證

5 個 Fold 之平均表現

Accuracy	91.80%
Precision	92.55%
Recall	91.18%
F1-score	91.08%



結果比較

	CNN	CNN + BiLSTM	RF1	RF2
模型種類	深度學習		傳統機器學習	
訓練時間	長		短	
運算資源	需求較高		需求較低	
使用特徵	訊號轉為頻譜圖		統計特徵	統計 + 頻率特徵
預測準確率	77%	86%	98%	92%



在目前之資料量及任務需求下，Random Forest 模型即可透過較低的時間與資源需求獲得更好的訓練成效

決策意涵

即時監測

使用 Random Forest 即可達到快速且有效的偵測（98%以上準確率），在機器發生故障前及時回報問題

節省人力

高預測能力將能降低原本需要定期派人檢查的人力成本

預測性
維護

在機器實際故障前可以及早偵測異常的機器，透過提前保養減少非預期故障所帶來的產能損失與維修費用

未來展望

收錄製造現場之資料

目前使用之訊號資料皆由機械故障模擬器產生，未來可收錄製造現場之感測器資料，確保模型能應用到真實情境

遷移學習與跨設備泛化測試

目前僅有偵測三種不同的零組件，可以嘗試偵測不同組件或是機器的聲音，將預測性維護機制拓展到整個工廠

自我監控與模型更新機制

可建構自動化監控模型效能機制，在模型偵測效果下降時自動觸發再訓練程序更新模型，維持長期準確度

應用情境

Demo

Real-Time Machine Audio Anomaly Detection

Monitoring folder: **sensor_data**

Drop sensor data files (CSV format) into the folder to detect anomalies.

No data processed yet.

Past Records

Show anomalies only

Timestamp

Anomaly Type

Confidence

Thank You

