

製造數據科學期末報告 第 12 組

組員

張寬新 B06303001 張智鈞 R09725051 熊才誠 B05504121 王捷弘 B05602038

壹、主題:

機器軸承預期壽命預測 (Reamain Useful Life Prediction of Bearing)

貳、背景與動機:

一、資料背景

資料總共分為 Learning 和 Test 兩資料集在 3 種不同 condition 下的軸承紀錄資料，並使用兩組 sensor 來分別記錄軸承的震動(水平及垂直)及溫度。其中 Learning_set 中的資料為記錄過去軸承運作至損壞的完成資料，Test_set 則是需要預測剩餘壽命的目標軸承資料。

資料的收集頻率在不同變數間並不相同，在溫度的部分是每 0.1s 記錄一次，並且穩定記錄。震動幅度則是每隔 0.0004s 記錄一次並記錄 2560 筆，然後每隔 10s 重複一次記錄的動作。

二、動機

軸承是多數機器中的重要零件，需要承受高強度的振動以及高溫，也因此，軸承往往因為過度的振動及高溫導致損壞，需要維修或是替代，造成工廠大量的維修成本以及人力的投入。

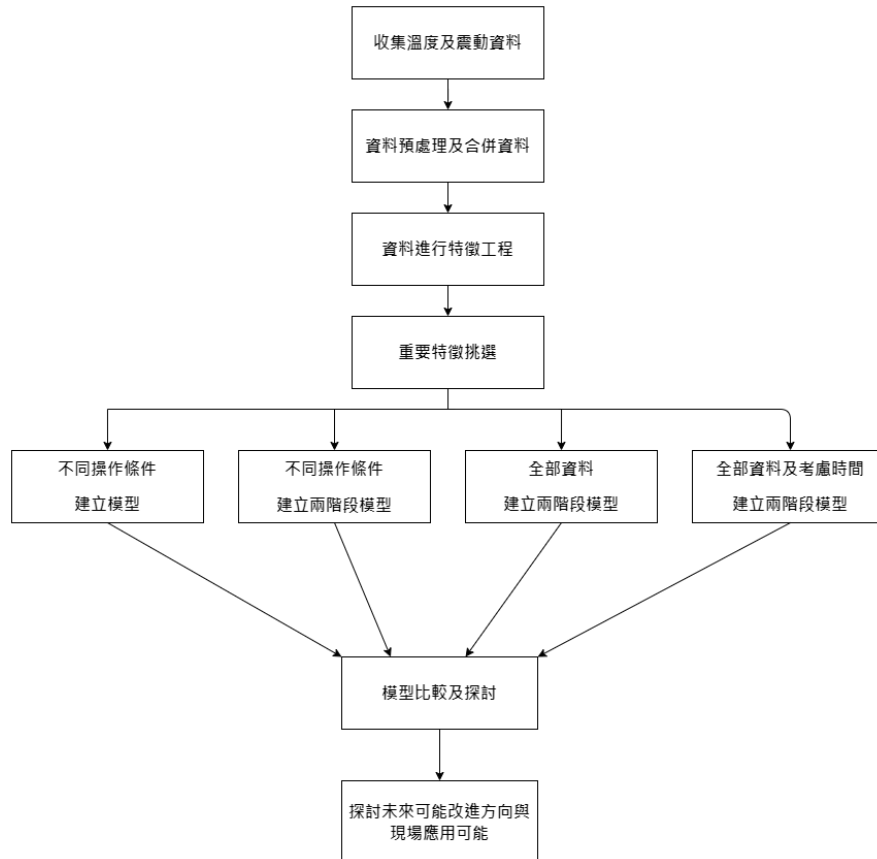
因此，藉由軸承使用壽命的預期，使得工廠能夠更精準的預測軸承剩餘的使用時間，藉此改善過去需等到機器故障才進行維修的狀況。在減少大量成本的同時，也讓產線更加穩定，如果未來再搭配上自動化的維修及檢查，將能省下大量的維修成本與人力資源。

三、問題定義

根據震動及溫度的變化來預測機器軸承的剩餘壽命

參、方法

一、研究框架



二、分析預測方式

利用不同的方法預測軸承剩餘壽命，我們透過比較不同的演算法 Linear Regression、Lasso Regression、Decision Tree Regressor、Random Forest Regressor、LightGBM 等五個模型的結果，來判斷哪一種模型會是最佳解，並且也考慮資料的合併或分割，加入時間考量，兩階段模型等等嘗試得到更好的軸承剩餘壽命預測。

肆、資料蒐集與分析結果

一、資料蒐集

資料集來源為 IEEE PHM 2012 Prognostic challenge 競賽的題目，題目是透過過去軸承的運作記錄，包括震動和溫度，來預測該軸承的剩餘使用壽命。

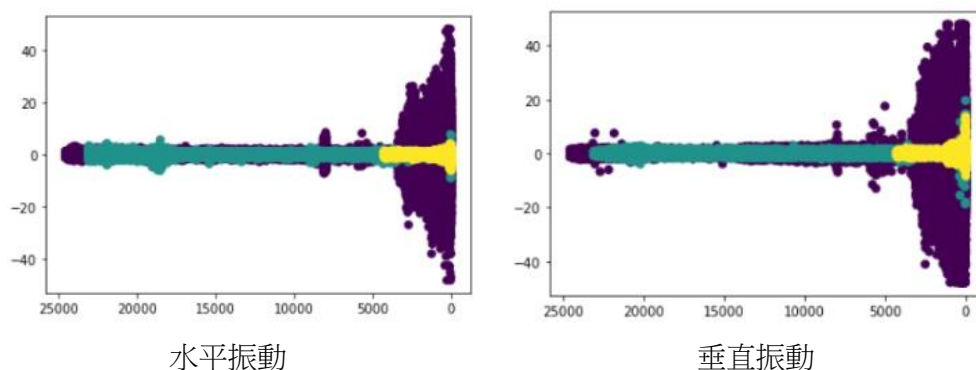
二、資料分析

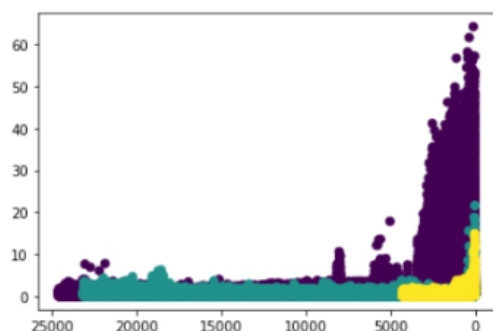
資料集內主要可分為兩種資料，一種是溫度，一種是震動幅度包括橫向和縱向，溫度的部分是每 0.1s 記錄一次，並且穩定記錄。震動幅度每 0.1s 記錄一次並記錄 2560 筆，然後每隔 10s 重複一次記錄的動作。而這種記錄方法的結果導致溫度和震動幅度的資料時間維度不同，因此我們假設溫度在 0.1s 的時間維度中是不會改變的，因此在同一秒內記錄的震動幅度資料假設其溫度相同，並在其資料集中加入溫度的 column，而在沒有記錄震動幅度時的溫度就直接捨棄。

而在某些資料集當中，只有出現震動幅度的記錄，在此報告中我們將其排除，因為我們想要將溫度和震動幅度兩種訊號都加入探討。然而如果我們將每個震動幅度資料都放入程式中預測，那就會導致資料量太龐大，因為每 10 秒就有 2560 筆資料，每個承軸記錄的總時長都超過三個小時，所以資料量總共會有千萬筆以上，因此我們將時間維度變大，也就是每 100 筆資料取一筆資料來加入預測，藉此來減少資料量。

接下來我們計算所有 data 中的 RUL，也就是利用最後一筆資料記錄的時間去扣掉中間每個記錄資料時間，這樣就把它當作是該記錄資料的 RUL。最後，因為震動方向包含橫向和縱向，也就代表他的震動方向是往斜的，因此我們再把橫向震幅的平方加上縱向震幅的平方再開根號，就變成斜向的震幅，也就是總震幅，並將其結果新增為一個 column。

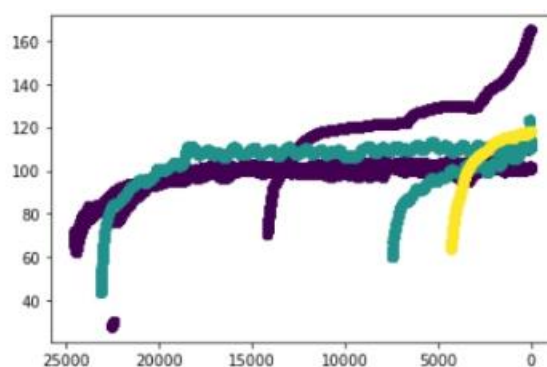
我們對讓新增的溫度以及總震幅的 column 對 RUL 做圖，下圖為總震幅-RUL 的圖，可以看出當 RUL 越小時總震幅會越大，代表越接近壞掉時承軸震動的幅度越大。





累積振動

而下圖為溫度- RUL 的圖，可以發現當 RUL 越小，溫度會越高，因此發現當承軸越接近壞掉的情況，震動所產生的溫度就越高。



溫度

(不同顏色是不同 condition)

三、資料預處理

而由於震動及溫度收集的頻率不同，震動的收集頻率更密集，我們利用填值的方式將兩種收集資料合一起，並且由於資料量太大，我們將震動改成每 0.1 秒收集 256 筆資料，並且將只有一種資料的實驗刪除。

四、特徵工程方法

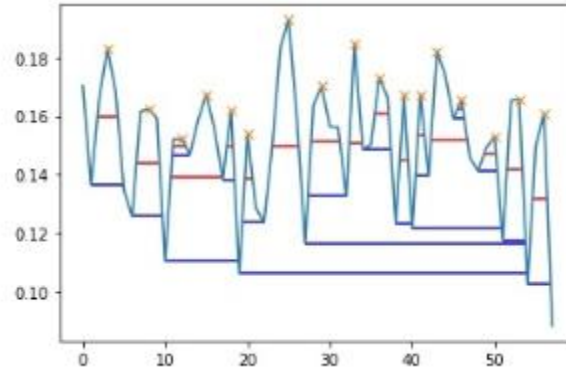
(一)時域相關特徵工程

時域是根據時間進行特徵工程，由於我們實際預測壽命可以知道過去幾個時間點的資料，所以去計算過去幾個時窗的值給出相對應的趨勢變化特徵，可能對於我們的預測更有幫助，以下介紹幾個我們用的特徵工程方法

1.移動平均 (Moving average)：計算前面幾個時窗不同感測器的值，計算他們的平均值成為最為新的特徵，可以讓我們的值變化的更為平滑有助於了解震動或是溫度隨著時間的變化趨勢。

2.最大值 (Peak Value)：算出前面幾個時窗的不同感測器的最大值，而這樣的離群值可能代表著故障情況，對於我們的壽命預測有幫助，

3.峰與峰的距離 (Peak to Peak)：去計算我們的震動頻率的變化，峰與峰的距離代表著我們的感測的波峰到下面一個波峰的距離，如果頻率較低則距離可能更大。



上圖為 Peak to Peak 示意圖，X 軸為時間，Y 軸為震動值的變化，而其中的藍線及紅線就是我們轉換出來的特徵值，其中我們也計算波的中間到下一個中間值的距離，如示意圖所示他們代表著不一樣的資訊，我們想要探討是否隨著時間的變化，這個距離也會有所不同。

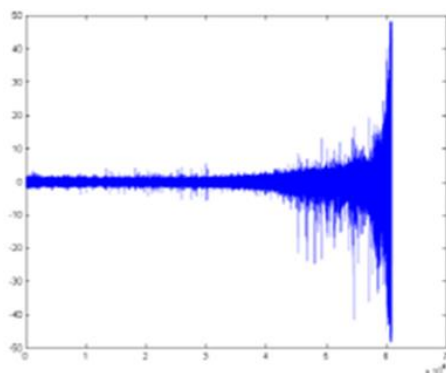
4.其他時域特徵工程：我們也參考了老師的講義去把更多的特徵值轉換出來，其中包含了變異數、偏度、峰度、峰質因數、波型參數等等，希望去探討我們的感測器更多的變化趨勢。

Mean	$u = \frac{1}{N} \sum_{j=1}^N n_j$
Standard deviation	$\sigma = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (n_j - u)^2}$
Root mean square	$\text{RMS} = \sqrt{\frac{1}{N} \sum_{j=1}^N (n_j)^2}$
Skewness	$\text{SK} = \frac{\sum_{j=1}^N (n_j - u)^3}{(N-1)\sigma^3}$
Kurtosis	$\text{KU} = \frac{\sum_{j=1}^N (n_j - u)^4}{(N-1)\sigma^4}$
Crest indicator	$\text{CI} = \frac{\max n }{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}$
Clearance indicator	$\text{CLI} = \frac{\max n }{\left((1/N) \sum_{j=1}^N \sqrt{ n_j } \right)^2}$
Shape indicator	$\text{SI} = \frac{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}{(1/N) \sum_{j=1}^N n_j }$
Impulse indicator	$\text{MI} = \frac{\max n }{(1/N) \sum_{j=1}^N n_j }$

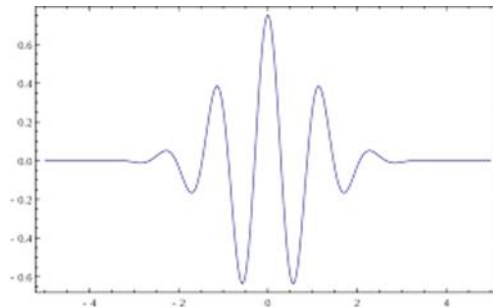
(二)時頻域領域特徵工程_小波轉換的應用

根據此競賽內容我們比對了其他參賽者所使用的方法，在頻率轉換的部分他們大多使用傅立葉轉換來解析頻率。然而，我們在根據資料觀察後認為傅立葉轉換的使用可能在此資料中有限制，因此選用了小波轉換，以下為我們選擇的原因及實際操作的結果。

由下圖可以觀察出資料前期波動振幅變化較不明顯，而後期波動振幅隨時間有加劇的現象。因此為了要取得各時段中的頻率組成及強弱資料(用以後續分析)，我們必須將原先的 Time-Domain 資料轉成 Time-Frequency Domain 的形式。在此，我們使用小波轉換來達成此目的。



(x 軸:時間, y 軸:振幅)

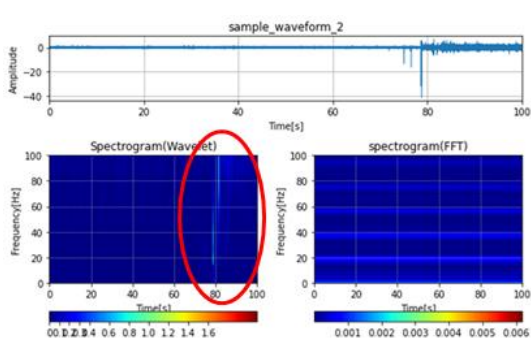
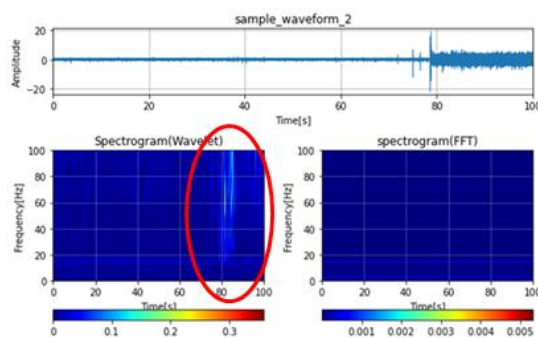


莫萊小波(morlet wavelet) 示意圖

小波轉換具有處理非平穩訊號的功能，並能將轉出的訊號也提供時頻表示，解決了傅立葉轉換無法處理受時間影響資料的缺點。下圖為小波轉換(使用 morlet 小波)後的結果(以 Bearing3-1 為例)，從圖中可以明顯看出不論在水平或是垂直方向上的振動，都在時間點 80 以後才有比較大的變化。而對照左下(小波轉換)與右下(傅立葉轉換)的圖，可發現透過小波轉換能在時間點 80 的位置精準的抓到高頻振動的出現，而傅立葉轉換則無法捕捉時間與頻率的關係。

垂直振動

水平振動



(圖上方為時間對振幅的結果,左下為小波轉換之結果,右下為傅立葉轉換之結果,顏色的深淺，代表該頻率的訊號強弱(與振幅成正比))

(三) 特徵工程參數選擇

時域的時窗大小透過物理特性決定，震動= 26 (0.01s)及溫度= 60 (6s)，並且溫度資料較沒有震幅的現象所以不做小波轉換及前面計算的 P2P，小波選擇的頻率從 8 到 256 並且每 8 個取一次，由於太接近的頻率可能代表類似的資訊，變成多餘的 feature 經過轉換我們的 feature 從原本的 4 個(溫度/水平/垂直/累積震動)，轉變成總共：154 個特徵值

五、特徵篩選方法

(一)篩選依據

經過 Feature engineering 後，我們產生了大量的變數，但變數過多可能會變成模型訓練上的阻礙，且許有許多變數代表的資訊相當類似，變得多餘，造成共線性的問題。因此，我們使用以下幾種 Feature Selction 的方式進行變數篩選，最後再利用投票法選出重要的 Features:

Correlation	比較變數與 RUL 之間相關係數的絕對值大小
Lasso Regression	將過於共線性的變數移除
Stepwise Regression	一步一步挑選變數值到達到最好的預測能力
Random Forest	利用建樹時常挑選到的變數
Light GBM	利用建樹時常挑選到的變數

(二)投票法依據及結果

我們使用上述的 5 種篩選變數方式各挑選前 50 個重要變數，如果變數同時在 2 個以上的方法中被挑中，則將其視為重要變數，以下為投票法產生之結果:

	Feature	Pearson	RFE	Regression	Random Forest	LightGBM	Total
0	rms_vert	True	False	True	True	True	4
1	rms_horiz	True	False	True	True	True	4
2	av_sum	True	False	True	True	True	4
3	Rtd sensor	True	False	True	True	True	4
4	sd_vert	True	False	True	False	True	3
...
142	clear_vert	False	False	False	False	False	0
143	clear_horiz	False	False	False	False	False	0
144	av_vert	False	False	False	False	False	0
145	av_horiz	False	False	False	False	False	0
146	Horiz_acc	False	False	False	False	False	0

(三)依據投票法在不同 setting 上的差異

由於不同 setting 對於 sensor 的影響性相當大，我們先決定採用分開建模的方式。而以下從三個 feature selection 的結果看出不同 setting 不同重要的特徵：

setting1		setting2		setting3	
Feature		Feature		Feature	
0	rms_vert	0	sd_vert	0	peak_temp
1	rms_horiz	1	rms_horiz	1	av_sum
2	av_sum	2	f16	2	Rtd sensor
3	Rtd sensor	3	av_sum	3	sd_vert
4	sd_vert	4	sd_temp	4	sd_temp
5	sd_temp	5	sd_sum	5	sd_sum
6	sd_horiz	6	sd_horiz	6	sd_horiz
7	rms_sum	7	rms_vert	7	rms_sum
8	peak_temp	8	rms_sum	8	rms_horiz
9	f96_vert	9	peak_temp	9	peak_sum

不同 setting 投票法前 10 名的重要變數

結果分析：

可以看出不同 condition 的資料集中，相對重要的特徵值也有所不同，幾個不同的地方是，condition 1 垂直及水平的震動的 RMS 較為重要，condition 2 則是垂直震動的變異數及經過水平震動經過小波轉換頻率為 16 的特徵，condition3 則是溫度最為重要。

六、模型建立

(一)模型挑選

經過特徵工程及特徵挑選後，我們把這些重要的特徵進入建模，並且由於我們前面的視覺化圖表可以看出，軸承不同的設定條件下，我們的感測器及預期壽命會有很大的不同，所以我們特過上面三個不同的設定條件，去訓練三個不同的模型，我們的資料集大小：

訓練集（2194175 x 52 / 782330 x 54 / 109568 x 46）

測試集（936447 x 52 / 230656 x 54 / 131838 x 46）

並且針對上面三個我們總共利用了以下幾個不同的模型：

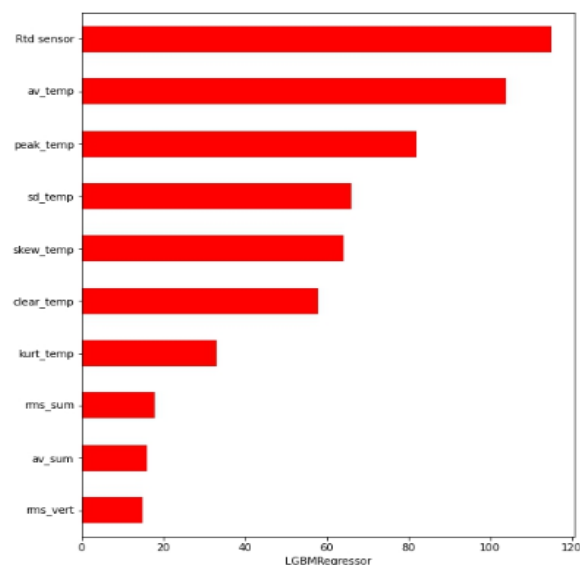
Linear Regression、Lasso Regression、Decision Tree Regressor、Random Forest Regressor、LightGBM

經過模型結果比較，我們發現 Random Forest 及 LightGBM 都有不錯的效果，而 LightGBM 模型有訓練時間快速且減少記憶體消耗的好處，更加適合我們這麼大量的資料，所以最終選擇 LightGBM 當成我們最終的模型。

(二)訓練結果

1. Setting 1 model

```
MSE training: 12953840.104, MSE testing: 70250852.179
RMSE training: 3599.144, RMSE testing: 8381.578
MAE training: 2740.729, MAE testing: 6596.583
```

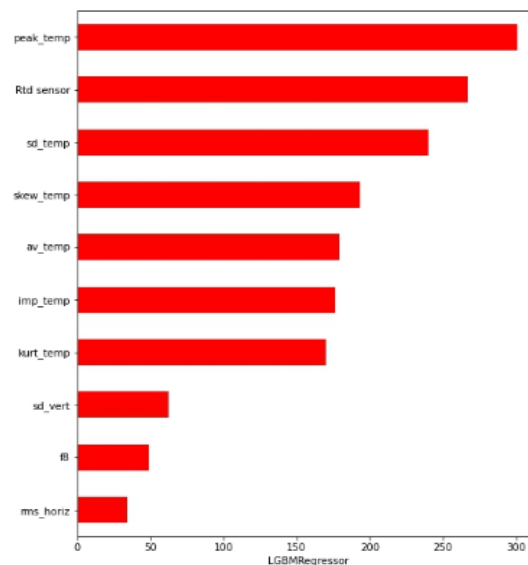


結果分析：

可以看出 testing 的 RMSE 及 MAE 較 training 高出許多，可能有 overfitting 的現象，並且參數重要性的比較中，溫度的值對於我們預測 RUL 最有效果，其次是累積震幅的 RMS 及平均數。

2. Setting 2 model

```
MSE training: 14020790.247, MSE testing: 190753735.036  
RMSE training: 3744.435, RMSE testing: 13811.363  
MAE training: 2870.178, MAE testing: 12927.630
```

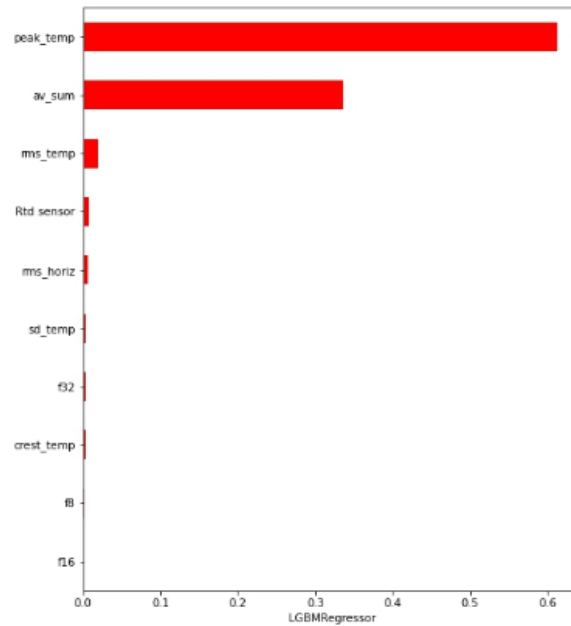


結果分析：

可以看出測試資料集中的 RMSE 及 MAE 較 training 高出許多，可能有 overfitting 的現象，並且參數重要性的比較中，溫度的最大值是最重要的變數，其次則是垂直震動的變異及水平震動經過小波轉換的特徵值(F=8)。

3. Setting 3 model

```
MSE training: 23570.222, MSE testing: 1798134.304  
RMSE training: 153.526, RMSE testing: 1340.945  
MAE training: 130.708, MAE testing: 1078.326
```



結果分析：

可以看出測試資料集中的 RMSE 及 MAE 較 training 高出許多，可能有 overfitting 的現象，並且參數重要性的比較中，溫度的最大值是最重要的變數，其次是平均累積震動且 setting 3 前兩名的變數貢獻度非常高其他就相對沒有那麼重要，可能模型建立大多數是依據這兩個變數。

七、兩階段模型

透過前面視覺化的圖可以發現，整個軸承損壞的過程感覺可以分成好幾個階段(穩定期/剛開始損壞/損壞造成劇烈變化)，不同時期有不同的特徵，我們希望將不同的時間分開建模，讓我們的模型預測能力提升。

而如何決定我們的資料應該分成幾個類別，我們透過 Grid Search 的方式去找群最正確的個數，比較不同個數 Testing 的 F1 score 找到最適合的個數，而由於前期我們的資料改變幅度較為不明顯，我們用 50% / 75% / 87.5% 的方式去切分資料，最後三個 setting 都是分成 3 類時 F1 score 最高。

(一) 分類模型

關於 Testing data，因為不能直接偷看答案來進行分類，首先需要先預測他的類別，再根據預測類別執行對應的 model，對應的 model 則是由該類別 Training data 訓練而產生的。

分類結果(LightGBM)：

Setting1	accuracy training: 0.655, accuracy testing: 0.777 F1 training: 0.612, F1 testing: 0.763
Setting2	accuracy training: 0.807, accuracy testing: 0.500 F1 training: 0.803, F1 testing: 0.333
Setting3	accuracy training: 0.998, accuracy testing: 0.500 F1 training: 0.998, F1 testing: 0.333

(二) 根據個別的模型預測預期壽命

最後根據我們預測的類別，去執行對應的模型，結果如下：

Setting1	MSE training: 30579216.755, MSE testing: 62960946.902 RMSE training: 5529.848, RMSE testing: 7934.793 MAE training: 3510.011, MAE testing: 6365.431
Setting2	MSE training: 14946360.270, MSE testing: 191003343.665 RMSE training: 3866.052, RMSE testing: 13820.396 MAE training: 2299.707, MAE testing: 12875.238
Setting3	MSE training: 3785.057, MSE testing: 1975315.478 RMSE training: 61.523, RMSE testing: 1405.459 MAE training: 49.545, MAE testing: 1111.956

結果分析：

我們發現相較我們前面直接用全部的資料去預測，此分段的模型有更好的效果，並且我們也可以不只去預測他的壽命，可以預測機器老化可能的階段，作為現場預警的功能。

八、模型再調整

由於訓練的結果不如預期，我們接著嘗試利用以下幾個方式進行模型的調整：

(一) Setting 分開與合併比較

我們發現前面的的模型不管是預測類別或是預測 RUL 都有 Overfitting 的現象，我們判斷是因為訓練資料實驗組數太少了，setting 2 只有兩組，setting 3 只有一組，造成模型很容易根據單一的實驗組數去預測，導致嚴重的 overfitting。

我們嘗試將不同的 setting 加再一起訓練並且重複 two step 的步驟，來增加實驗組數也探討實驗組數與 setting 影響兩者何者較為嚴重。

新模型結果：

全體 setting	MSE training: 22361511.219, MSE testing: 73191425.218 RMSE training: 4728.796, RMSE testing: 8555.199 MAE training: 2978.731, MAE testing: 6391.750
setting 1	MSE training: 29225119.740, MSE testing: 55382681.082 RMSE training: 5406.026, RMSE testing: 7441.954 MAE training: 3800.457, MAE testing: 5803.921
setting 2	MSE training: 19733894.855, MSE testing: 81877031.185 RMSE training: 4442.285, RMSE testing: 9048.593 MAE training: 2664.808, MAE testing: 6790.384
setting3	MSE training: 25974274.803, MSE testing: 42654584.829 RMSE training: 5096.496, RMSE testing: 6531.048 MAE training: 3398.030, MAE testing: 4588.675

結果分析：

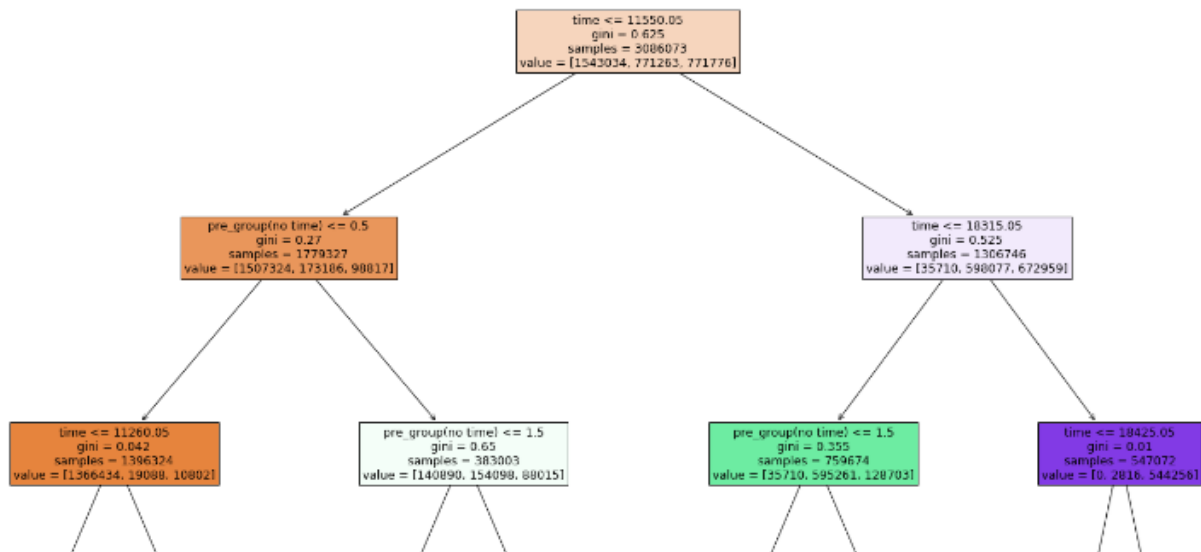
相較本來的兩階段模型，如果我們利用全部的資料去建模再 setting2 和 setting3 中有較好的預測結果，可能由於 setting 2 和 setting 3 的資料集較少，而相對再 setting 1，我們的訓練結果就較差，可能代表本來 setting 1 的資料就比較充足，全部資料去建模反而會受到 setting 的影響降低預測能力。

(二) 加入時間的考量

如果真的實際使用，那機台運作的時間其實是可以取得的，應該可以加入模型中，不過由於我們資料組數過少，當加入時間的變數，會造成嚴重的 overfitting，模型會大量挑選實際時間這個變數，並且推算出實際訓練資料的 RUL。

我們嘗試利用 Ensemble 的方式，再 Two Step 前段預測類別時，將預測的 類別與資料實際時間當成 X，再去訓練簡單的模型 (Decision Tree) 預測類別，希望可以減少 Overfitting 的發生也希望考量實際時間可以加強我們的模型，更能符合現場應用。

Decision Tree 樹狀圖:



模型分類老化階段結果:

accuracy training: 0.904, accuracy testing: 0.697
F1 training: 0.904, F1 testing: 0.698

模型預測結果:

全體 setting	MSE training: 9949930.070, MSE testing: 82600848.849 RMSE training: 3154.351, RMSE testing: 9088.501 MAE training: 1863.206, MAE testing: 7454.049
setting 1	MSE training: 4682547.256, MSE testing: 91617890.192 RMSE training: 2163.919, RMSE testing: 9571.723 MAE training: 1392.753, MAE testing: 7992.105
setting 2	MSE training: 20919498.435, MSE testing: 59301590.659 RMSE training: 4573.784, RMSE testing: 7700.753 MAE training: 2815.143, MAE testing: 6103.827
setting3	MSE training: 37108747.110, MSE testing: 59315698.758 RMSE training: 6091.695, RMSE testing: 7701.669 MAE training: 4487.398, MAE testing: 5994.505

結果分析：

從樹狀圖可以看出分類樹利用時間加上我們預測的 class，同時可以考慮這兩個東西去判斷我們機器的老化階段可以更準確，現場也能透過簡單的 if - else 準則去做判斷，增加此模型的可用性，而在預期壽命的預測上，我們發現加入時間沒有得到更好的結果，而是更加 overfitting，很主要的原因是因為我們的實驗組數很少，所以是直接利用 training 來做 stacking，未來可能還可以進行調整。

伍、結論

一、未來改善方向

(一) 此資料集數量相當多但實驗組數較少，造成後續的訓練結果不理想，如果可以的話應該再納入一開始沒有溫度資料的實驗組，增加資料組數，但經過分析後溫度的資料其實相當重要，可能未來可以利用 ensemble 或是填值的方式來增加溫度的數據

(二) 不同 Setting 的結果差異很大，未來可能可以利用其他技巧消除或正規畫其影響，使得模型可以訓練同一個增加預測力，或是將 Setting 的數值加入模型當中

(三) 同個 Setting 中 RUL 的變異性相當大，判斷可能由於其他因素去影響機台的運作時間，未來可以增加 Sensor 收集的種類更能探討實際情況，或是收集更多實驗組更了解整體 Setting 的實際影響

二、結果應用

(一) 透過 sensor 的收集，使用我們的模型可以即時地預測機台的壽命，並及時保養。藉此能有效改善過去需等到機器出現嚴重故障才進行維修的狀況，減少大量的成本，也讓產線更加穩定

(二) 在過去進行機台狀況的評估中，往往只能夠過一些物理/機械的知識輔助，而這些知識的假設往往與現場有些微落差。而現在我們能透過 sensor 及大數據分析更能了解機台的狀況，並且可以透過重要特徵挑選，對應 sensor 收集機器的物理特性，有更新的 insight，更能輔助往後的機台保養與機台設計

(三) 未來機台保養有望朝自動化方向發展，除了預測壽命以外，也預測哪部分的零件有損壞，並且自動化的進行維修，節省大量的成本及人力

陸、參考文獻

(1) A Model-Based Method for Remaining Useful Life Prediction of Machinery Written By Yaguo Lei, Member, IEEE, Naipeng Li, Szymon Gontarz, Jing Lin, Stanislaw Radkowski, and Jacek Dybala 2016

(2) A Two-Stage Approach for the RUL Prediction of Bearings Using Deep Neural Networks Written By Min Xia , Teng Li , Tong xin Shu , Jia fu Wan , Clarence W. d 2019