|  |  |
| --- | --- |
| 大專新進人員訓練工作報告  　姓　　名： 王翔禾 單位：　台化工務部自控處  　報告類別：□輪班訓練　 ■基層實務工作訓練第 (三) 次報告  　訓練部門：台化工務部自控處智慧專案組  　起迄日期：2023/06/14~2023/09/14    工作項目：ARO2廠甲苯塔優化模組人機介面建置及營建鋼構最優化設  計AI模組開發  　報告項目：ARO2廠甲苯塔優化模組人機介面建置及營建鋼構最優化設  計AI模組開發 | |
| 評 核 ( 評 語 ) | |
| (２) 一 級 主 管 | (１) 二 級 主 管 |
|  |  |
| (４) 經 營 主 管 | (３) 經 理 室 |
|  |  |

表號：P0002703 規格：A4

註：一.報告內容應包括訓練期間之 1.工作心得感想 2.所得之工作觀

念及精神 3.自我檢討 4.建議意見。

二.訓練部門主管應詳細評核「訓練工作報告」內容，並批註意見。

**目錄**

[一、 前言 1](#_Toc141869725)

[二、 ARO2廠甲苯塔優化模組人機介面建置 4](#_Toc141869726)

[2-1系統架構 4](#_Toc141869727)

[2-2系統開發流程 7](#_Toc141869728)

[三、 鋼構最優化設計AI模組開發 10](#_Toc141869729)

[3-1鋼構最優化設計AI模組開發數據清理 11](#_Toc141869730)

[3-2鋼構最優化設計-數據模型建立與評估 15](#_Toc141869731)

[3-3鋼構最優化設計-基因演算法開發 17](#_Toc141869732)

[3-3-1鋼構最優化設計-基因演算法說明 20](#_Toc141869733)

[3-4鋼構最優化設計-結果呈現 23](#_Toc141869734)

[3-4-1鋼構最優化設計-族群數量比較 24](#_Toc141869735)

[3-4-2鋼構最優化設計-演化代數比較 27](#_Toc141869736)

[四、 工作心得 30](#_Toc141869737)

前言

職目前承辦工作為「ARO2廠甲苯塔優化模組人機介面建置」及「鋼構最優化設計AI模組開發」。

第一案「ARO2廠甲苯塔優化模組人機介面建置」此案旨在協助各部針對單元製程建置可供單元操作調整之可視化工具，以網頁的方式呈現，即時提供節能操作條件建議，以維持最佳能耗操作並貼近品質管制基準。

甲苯塔C670在操作上僅依檢驗數據被動調整，且為避免品質超限(塔頂A8<2.0%，塔底Tol<1.0%)，採取保守操作，尚有節能空間，此製程於實務操作上，為避免超過管制基準，製程採取保守操作，因而有過度加熱現象，因此須建立節能優化AI模組，並架設Dashboard 以互動式網頁之方式提供現場人員操作判斷依據。。

第二案為「鋼構最優化設計AI模組開發」，鋼結構設計一般是設計者採用try and error(試誤法)方式，利用結構分析程式(如SAP2000)反覆調整桿件尺寸重複執行分析，直至所有桿件皆可符合規範要求(如:所有桿件之P-M Interaction Ratio 必須小於1)，除無法確保是否為最優化的斷面，過程亦需耗費大量工時，因此本案配合營建設計處，將專業結構分析程式(SAP2000)結合已被廣泛應用於解決最佳化和優化問題之AI演算法，開發內嵌優化設計演算法之程式，自動反覆執行結構分析、更換優化斷面等自動程序運作，取代傳統的人工試誤法，達成鋼結構設計最優化(最經濟)之目標、減少碳排及節省反覆執行程式所需時間之目標。

本案的AI模型可分為兩個主要部分。首先，利用大數據建立一個機器學習AI模組，用於預測鋼構架的初始桿件排列。這個模組能夠通過分析大量數據和學習結構特徵該案場的預測初始桿件排列。

其次，我們採用基因演算法作為第二部分，將第一部分預測的初始桿件排列作為基因演算法的一個初始化條件(初始化個體)。基因演算法是一種自適應搜索算法，通過模擬自然選擇和遺傳機制來找到問題的最優解。通過這個算法，我們能夠自動反覆執行結構分析、更換斷面等一系列操作，以代替傳統的人工試誤方法。在這個過程中，系統能夠逐步發現並優化設計，使得最終的結構設計更加優越和高效。

為了實現這一優化過程，我們採用了自行開發的API程式，該程式能夠實現自動化的結構分析和操作。通過使用這個自動化過程，我們能夠節省大量的時間和人力資源，並且找出最佳的設計方案，以滿足項目的需求和要求。

總結而言，本案的AI模型結合了機器學習和基因演算法的優勢，通過自動化的分析和優化過程，為鋼構的設計提供了高度精確和高效的解決方案，取代了傳統的人工試誤法，讓開發能夠更準確迅速。

本季工作心得，以上述二案為主軸，詳如後續說明。

1. ARO2廠甲苯塔優化模組人機介面建置

此案主要將ARO2廠建模完成的AI優化模組，依據AI模組所需即時數據及操作推薦值以互動式網頁的方式作為人機介面呈現，用戶端可以網頁方式查看操作推薦值(於管制基準下之推薦溫度、流量等)、即時數據(如進料流量、進料組成和塔頂、塔底溫度等)，並將即時數據處理後儲存於客戶指定的廠區品管資料庫，設計網頁程式將常用的參考數據從資料庫擷取出來以趨勢圖的方式呈現。同時網頁畫面也建置輸入欄位，可供使用者輸入各塔頂、塔頂底管制成分所期望之品質管制試算值，並以AI優化模組預測出試算的溫度流量等，以即時且直觀的方式輔助現場人員操作判斷。利益。

2-1系統架構

以下為本案工程開發時所用到的作業系統、資料庫、使用的程式語言，如表2-1所示。系統架構圖如圖2-1所示，使用Python語言之第三方框架「Django」開發連接企業網路的網頁後端程式及讀取PI資料庫中的數據，接續以Python語言設計將下載之數據轉存至化一部專用的PostgreSQL品管資料庫，後續再以 JavaScript等前端語言，將資料庫數據及模型預測結果，以網頁之形式架設於Apache Server中，以供使用者查看。

|  |  |
| --- | --- |
| 作業系統 | Windows 10 |
| 資料庫 | PostgreSQL、PI System |
| 程式語言、框架 | Python、JavaScript、Django |
| 網頁伺服器 | Apache Server |

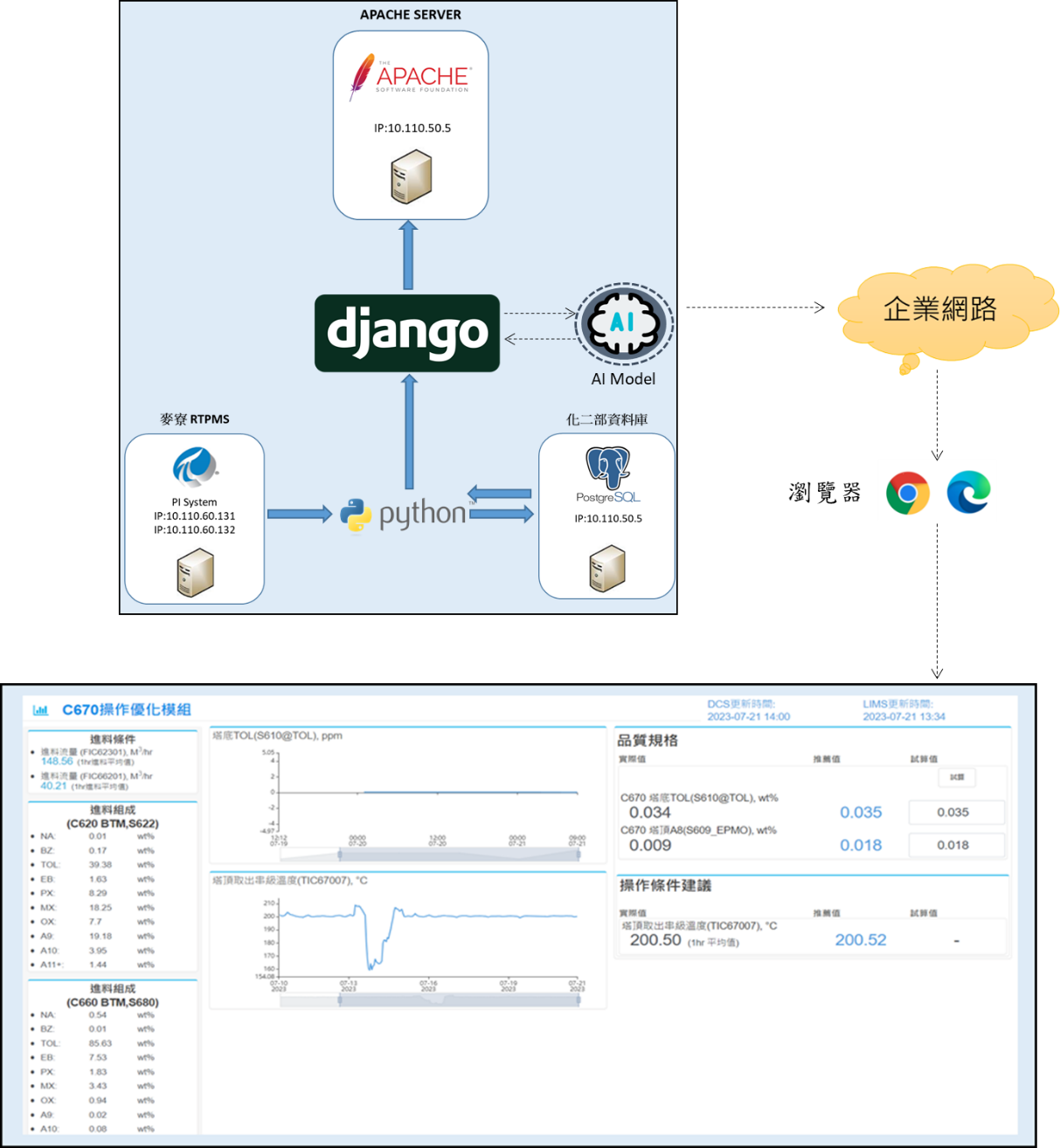
表2-1、系統環境

圖2-1、系統架構圖

2-2系統開發流程

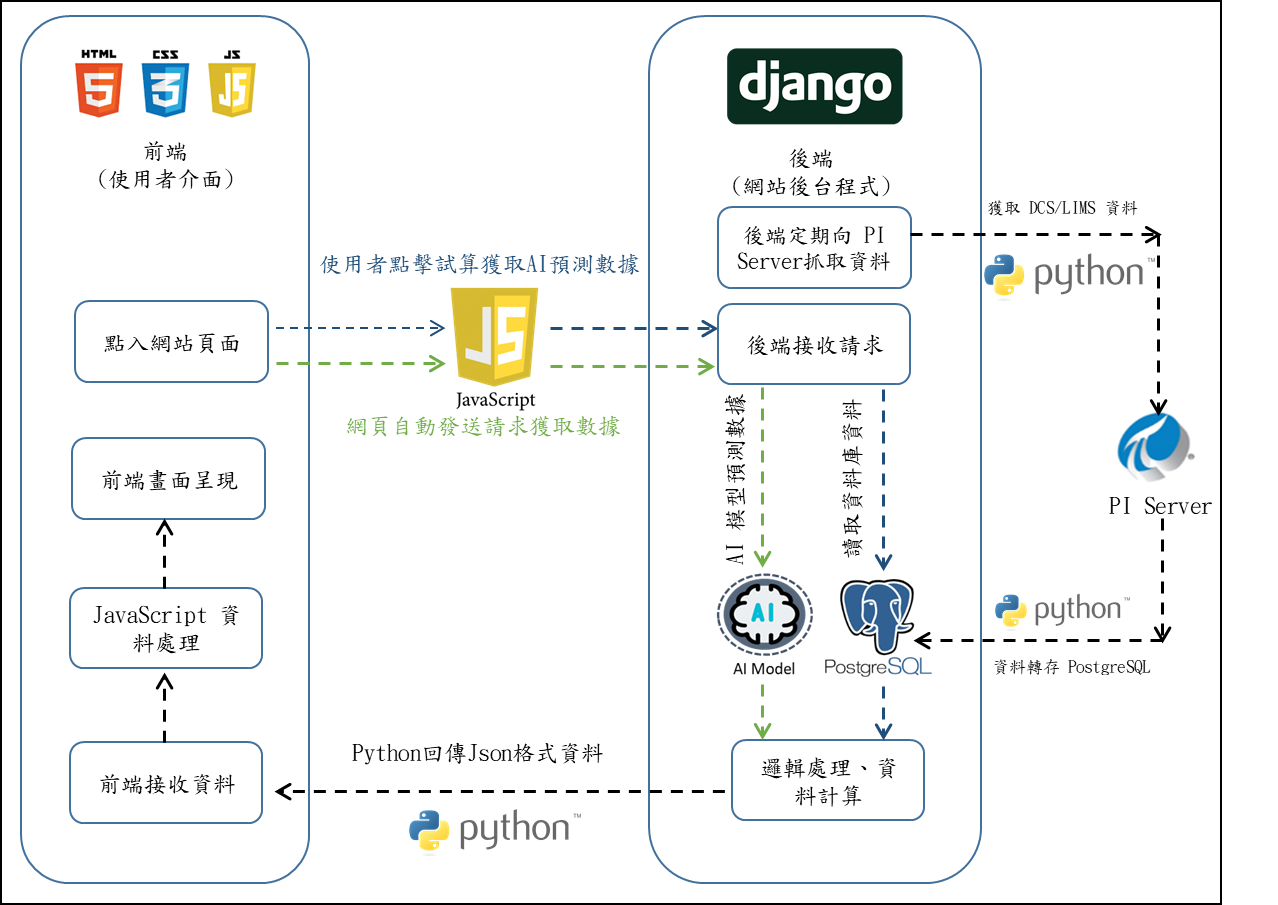
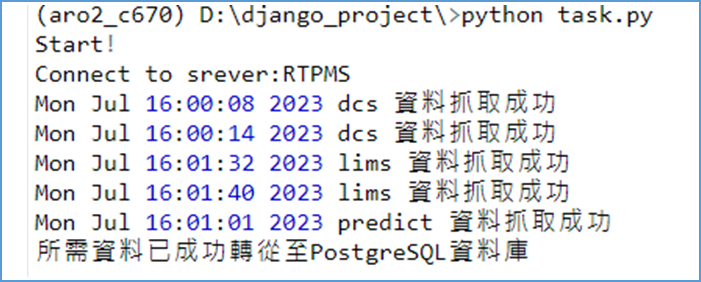
程式設計流程如圖2-2所示，於瀏覽器輸入網址後，前端JavaScript程式會發送資料請求傳遞至後端，後端則透過Django、Python開發之程式讀取PostgreSQL資料庫的資料，接著將讀取資料進行整理後，將資料以JSON型態回傳至前端，最後經JavaScript處理完後呈現使用者畫面，而後端 Python 程式也會定期向 PI Server 抓取 DCS/LIMS 數據，設定排程為每小時抓取最新資料，經處理後轉存至PostgreSQL資料庫以確保前端呈現之數據均在最新狀態，資料抓取程式畫面如圖2-3所示。

圖2-2、程式設計流程

圖2-3、定時資料抓取式畫面

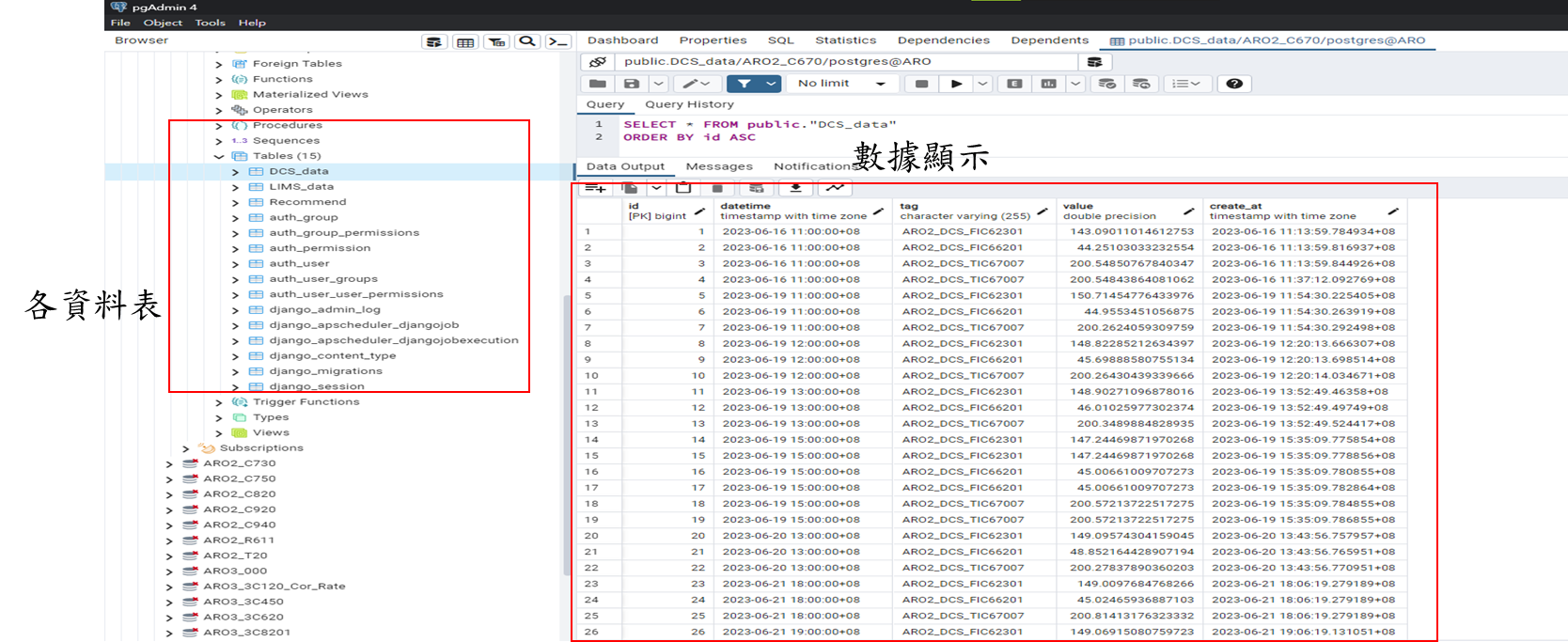
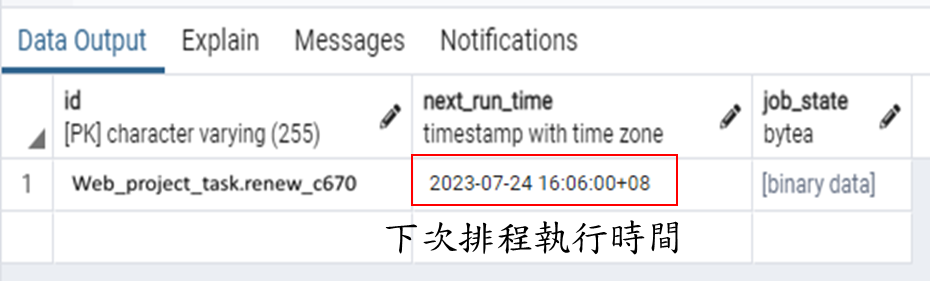
DCS資料彙總到PostgreSQL資料庫後可藉由PostgreSQL自帶的UI介面pgAdmin 查看資料庫內容，如圖2-4 、2-5所示，開發人員可隨時於資料庫中查看所需資料。

圖2-4、PostgreSQL資料表示意圖



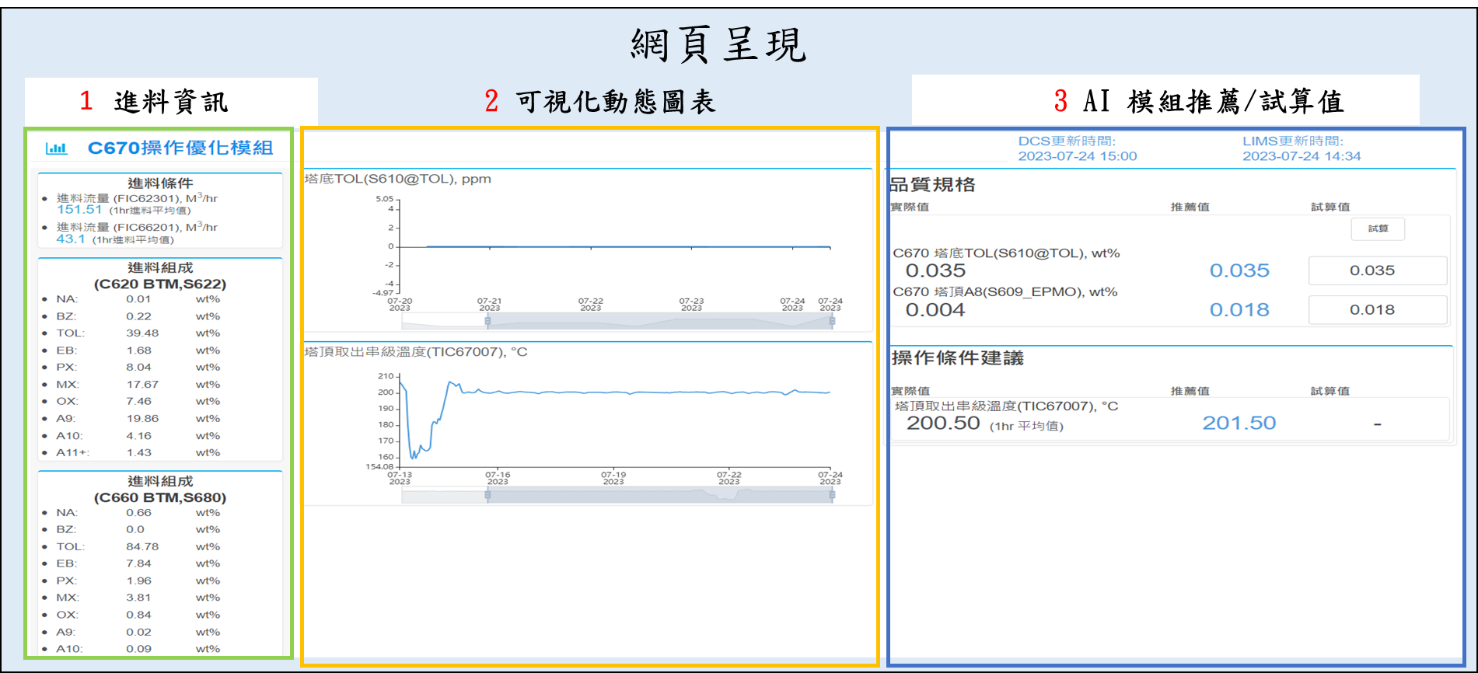
圖2-5、資料庫定時資料抓排程預計時間

圖2-6、AI人機介面互動式網頁呈現

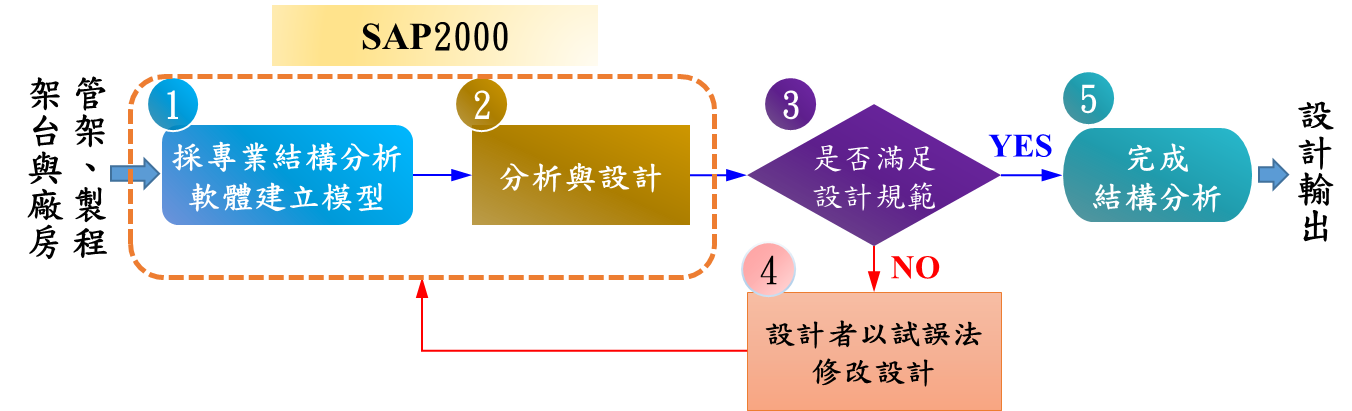
1. 鋼構最優化設計AI模組開發

在鋼結構設計過程中，設計人員必須依賴試誤方法，搭配其專業知識以及專業分析軟體SAP2000，耗費大量工時來探尋最經濟的桿件組合。

為改善此問題，擬開發AI模型來探尋最經濟的桿件組合，在與營建設計處討論後，將建置兩個模型分別為利用歷史數據配合機器學習得到能預測初始化桿件之「數據模型」，再將數據模型預測的初始桿件作為第二個模型「基因演算法型」的初始化條件之一，並透過自適應搜索算法，模擬自然選擇和遺傳機制來找到問題的最優解。

在人工智慧學校受訓期間，以鋼構最優化設計AI模組之「數據模型」開發作為專題，建立一個融合機器學習模型與專業軟體SAP2000驗證的解決方案，並搭配後續開發之基因演算法與自動化程式，以節省試誤法所需的時間與資源，圖3-1為營建設計流程。

透過AI模型的建立，我們能夠以更有效率且智能的方式來進行鋼結構設計。AI模型將透過大數據的分析與機器學習技術，學習過往的設計資料，並預測最佳的桿件組合。而此預測將成為SAP2000分析的起始條件之一，透過自行開發的API程式，自動進行結構分析、斷面更換等流程，以找出最優化的結構設計。



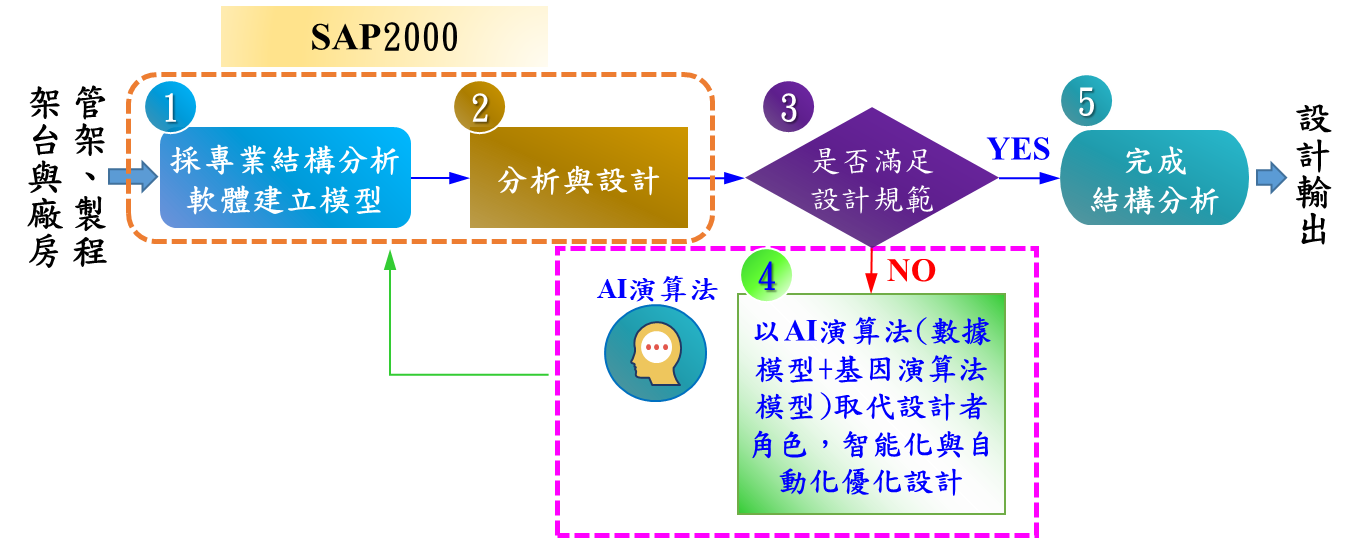
圖3-1、鋼構設計流程圖

圖3-2、鋼構最佳化設計流程圖

3-1鋼構最優化設計AI模組開發數據清理

本專題為一小組共五人合作進行開發，職在這案中主要負責異常資料處理、資料視覺化及機器學習建模，下方將對整案進行介紹。

根據營建設計處提供的資料，涵蓋27個案場，包含14組管架鋼構與13組鋼構廠房，皆為ARO3廠興建的設計廠房。總體數據共有28,735筆，每筆包含115個變數。

在與營建設計處進行深入討論後，便對原始資料進行特徵篩選，以確定對預測目標具有最重要影響的變數。同時，我們針對某些變數進行了特徵變換的處理，以更好地反映其在預測模型中的影響。

最終，經過篩選和變換，我們選取了14個特徵變數，，以形成最終的特徵集。這些特徵變數將與我們的預測目標一起用於建立預測模型，旨在準確預測相關設計廠房的桿件組合。

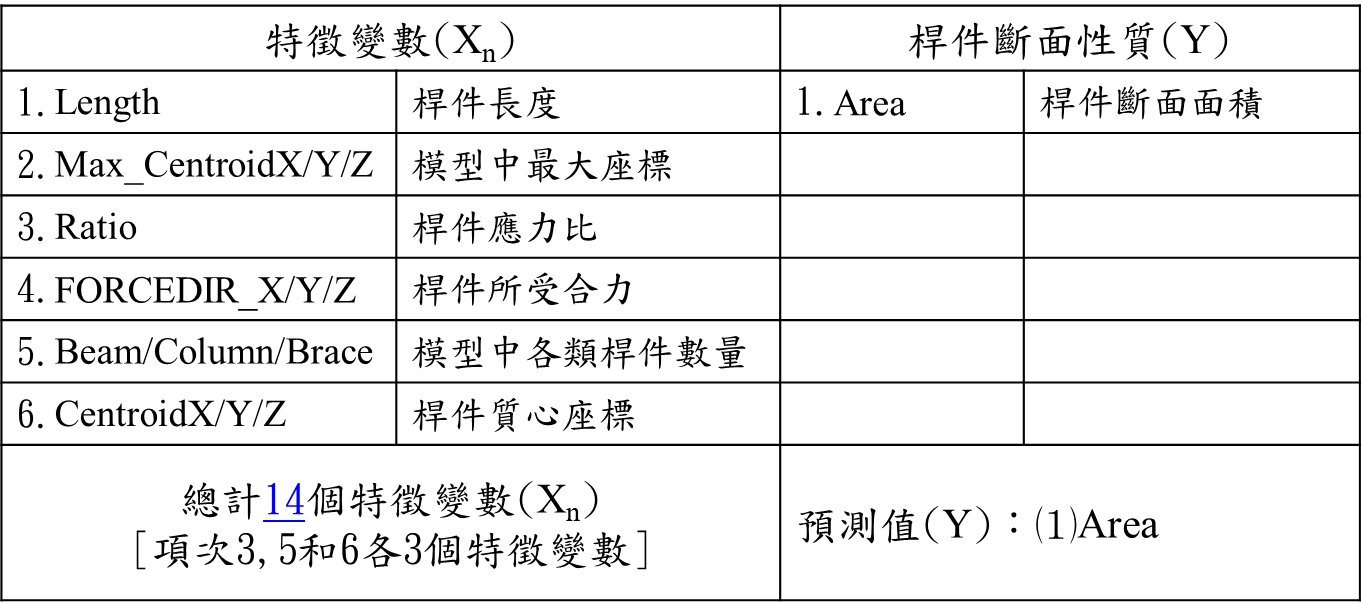


表3-1、鋼構設計特徵變數及目標



圖3-3、鋼構最佳化設計數據模型建模流程

經過特徵篩選後，對資料進行初步分析。首先，我們排除了含有空值的樣本，以確保資料的完整性和可靠性。其次，我們對數值型特徵進行了離群值的處理，採用三倍標準差的方法進行判斷，將超出此範圍的離群值排除。

這項處理過程對資料的品質提供了保障，共有985筆資料因含有空值或被視為離群值而被從原始資料集中排除。剩餘的27,750筆資料經過篩選後，將成為我們進一步進行建模和分析的有效樣本。

這樣的資料處理過程是為了確保分析結果的準確性和可靠性，同時排除可能影響模型效能的干擾因素。

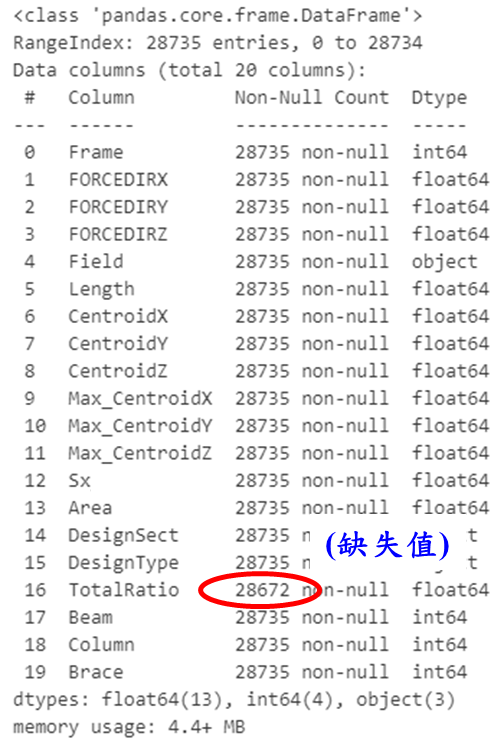
圖3-4、特徵變數離群值

圖3-5、特徵變數缺失值

經過資料處理後，繼續進行皮爾森相關係數的相關性分析。原本計劃篩選出相關性大於0.4的特徵值，結果共有9個特徵值符合此條件，顯示這些特徵與預測目標「鋼構斷面面積」之間存在較強的線性相關性。

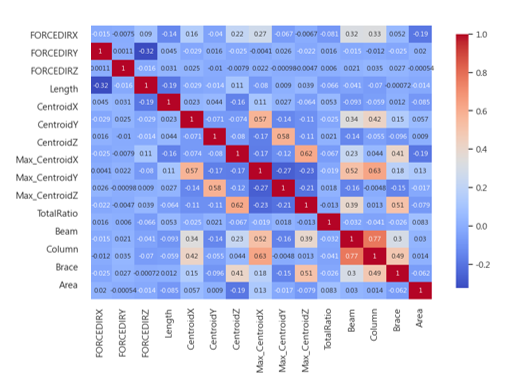
然而，到其餘五個特徵變數仍然保留一定程度的相關性，雖然未達到0.4的閾值。但在經過與設計人員的深入討論後，雖然它們的相關性未達到預定的閾值，仍然可能對於預測目標「鋼構斷面面積」有一定的影響決定保留這14個特徵變數進行後續的建模工作，下圖3-6為皮爾森相關係數分析圖。

圖3-6、皮爾森相關係數分析

3-2鋼構最優化設計-數據模型建立與評估

經過以上的前處理，我們的數據集包含了27,750筆資料。為了進行預測目標「鋼構斷面面積」的模型訓練，我們將這些資料分為訓練集和測試集，其中訓練集包含24,975筆資料，而測試集則包含2,775筆資料。

在進行模型訓練時，我們採用了四種不同的方法，分別是XGBoost、LightGBM、Random Forest和Support Vector Machine（SVM）。這些都是廣泛應用於機器學習中的強大算法，且都適用於數據量不多之訓練集，因此在本案中被選為訓練模型的方法。

在進行模型訓練時，我們將訓練集的數據提供給這四種方法，讓它們學習數據的特徵和目標之間的關係。接著，我們使用測試集來評估這些模型的表現，測試它們對「鋼構斷面面積」的預測準確度，其測試結果如下表3-2。

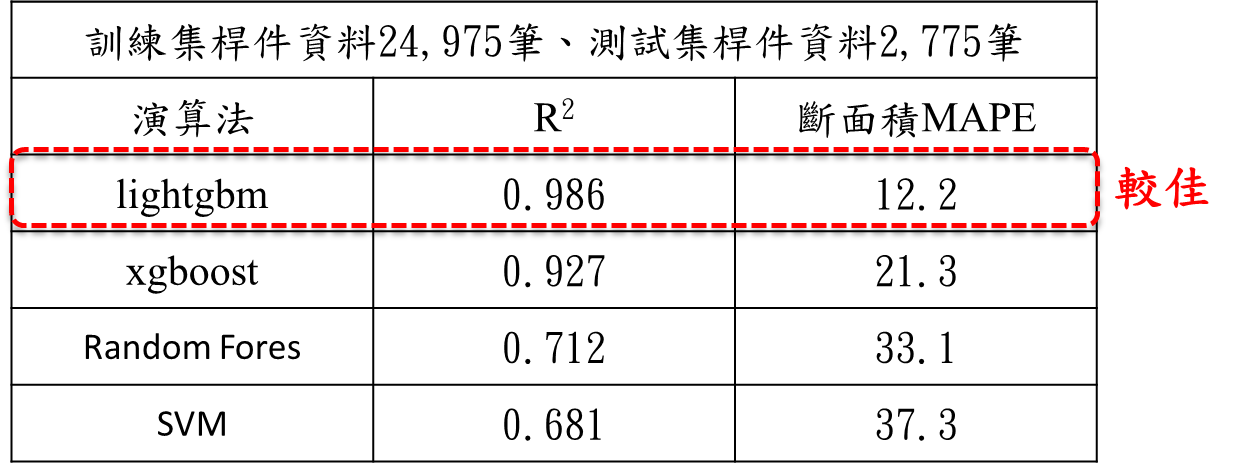


表3-2、太陽能數據監視系統軟體架構圖

經過對四種演算法（LightGBM、XGBoost、Random Forest和SVM）在訓練集和測試集上的評估，我們獲得了它們在預測目標「鋼構斷面面積」上的表現指標，包括R2和斷面積MAPE。

在R2評估指標中，LightGBM呈現出極佳的預測能力，達到了0.986的高度準確性，這意味著模型能夠很好地解釋目標變數的變異。相比之下，其他演算法的R2分數分別為0.927（XGBoost）、0.712（Random Forest）和0.681（SVM），顯示出LightGBM在模型擬合和預測方面優勢明顯。

此外，在斷面積MAPE（Mean Absolute Percentage Error）方面，LightGBM同樣表現出色，只有12.2%的預測誤差，顯示其對於預測目標的準確度較高。而其他演算法的斷面積MAPE分別為21.3%（XGBoost）、33.1%（Random Forest）和37.3%（SVM），相對於LightGBM，有較大的預測誤差。

綜合以上指標的比較，我們得出結論：LightGBM模型在本次預測任務中表現最佳，其高準確性和較低的預測誤差使其成為最適合的選擇。因此，我們選擇採用LightGBM作為最終的預測模型，並將其預測結果導入下一步基因演算法作為初始值(初始族群的其中一個個體)，以便準確地預測鋼構斷面面積。

3-3鋼構最優化設計-基因演算法開發

在完成3-2章節的數據模型建立後，我們將使用數據模型預測出的最佳初始桿件組合作為「基因演算法」模型的初始條件。基因演算法型模型將借鑒自然界的適應性搜索和遺傳機制的思想，透過自適應搜索算法來尋找問題的最優解。這個過程中，模型將持續進行交叉、變異和篩選等操作，以逐步優化桿件組合，並逼近全局最優解。

這兩個模型的協同運作將充分發揮歷史數據和演算法的優勢，為我們提供高效且精確的桿件設計結果。數據模型為基因演算法型模型提供了一個初始參考點，這有助於縮小基因演算法模型的搜索空間，從而加速求解過程。而基因演算法型模型則通過演化機制不斷優化解空間，以找到最優解。

透過這種結合了數據分析和演算法的AI模型，我們能夠更快速、更準確地找到最佳的桿件組合方案，從而優化建築設計過程並提升設計品質。

**3-3-1鋼構最優化設計-API開發**

在進行基因演算法的開發之前，為了能夠有效進行模型的交叉、變異和篩選等操作，我們必須先開發能與SAP2000對接的API程式。

由於基因演算法型模型需要不斷優化桿件組合並尋找最優解，這將涉及大量的運算和結構分析。因此，我們需要開發一個API程式，以便與SAP2000軟體進行無縫對接。這個API程式將允許我們以自動化的方式向SAP2000傳遞參數、模擬結構性能，並取得計算結果，從而實現基因演算法在實際結構分析中的應用。

透過API程式的開發，我們能夠高效地執行多次SAP2000的運算和分析，進而加速基因演算法的求解過程。這樣的整合將為我們提供更快速、準確的結構性能預測，同時減少了手動操作的時間和人力成本。

開發API程式是一項具有挑戰性的任務，需要充分理解SAP2000軟體的功能和運作原理，並與其相互配合，才能利用SAP2000軟體的強大功能，並將其應用於基因演算法模型的優化過程中，下圖3-7、圖3-8為API功能示意圖。



圖3-7 SAP2000 API 啟動程式



圖3-8 SAP2000 API 結構分析程式

3-3-1鋼構最優化設計-基因演算法說明

基因演算法（Genetic Algorithm，GA）是一種模仿自然界進化過程的優化算法，用於解決複雜的優化問題。其基本思想是通過模擬自然界的遺傳、交叉和適應性篩選等過程，逐漸優化候選解，以尋找問題的最優解。

以下是基因演算法的基本流程：

1.初始化：隨機生成初始個體群體，每一個體代表一個解， 並賦予它們隨機的基因組合。

2.評估適應性：對每個個體進行評估，計算它們的適應度（Fitness），適應度反映了個體在解空間中的優劣程度。適應度的計算與問題的目標函數相關。

3.選擇：根據適應度選擇一定數量的個體作為下一代的父代，通常適應度越高的個體被選中的機會越大。

4.交配：將選中的父代進行交叉操作，產生新的子代個體。交配的目的是模擬基因的交換，以產生新的個體組合。

5.突變：對子代進行變異操作，通常是隨機改變某些基因的值。變異的目的是保持種群的多樣性，避免陷入局部最優解。

6.終止條件：根據預設的停止條件（如達到一定的迭代次數或適應度達到一定閾值），判斷是否終止演算法。

7.返回第2步：如果未滿足終止條件，則回到第2步繼續進行演算法。

將基因演算法應用於本專案中，我們需要將桿件組合作為個體，每個個體的基因組合代表一種桿件配置方案。在初始化階段，隨機生成桿件組合，並搭配數據模預測出的初始化桿件當作基因演算法的初始族群。接著，進行評估步驟，計算每個桿件組合的適應度，這裡的適應度可以根據設計目標（如結構穩定性、成本等）的優劣程度來評估。

在選擇步驟中，根據適應度選擇一部分優秀的桿件組合作為父代，然後進行交叉和變異操作來產生子代。子代繼承了父代的優點，同時通過交配和變異的操作，保持種群的多樣性，並重複以上步驟進行多代的演化。

透過這樣的迭代過程，基因演算法能夠在大量解空間中搜索，並逐漸優化桿件組合，最終找到優秀的設計解。藉由基因演算法的適應性搜索能力，我們能夠探索更多可能的桿件組合，並在多個目標間找到平衡，達到最優化的設計方案。

基因演算法流程圖，如下圖3-9所示。

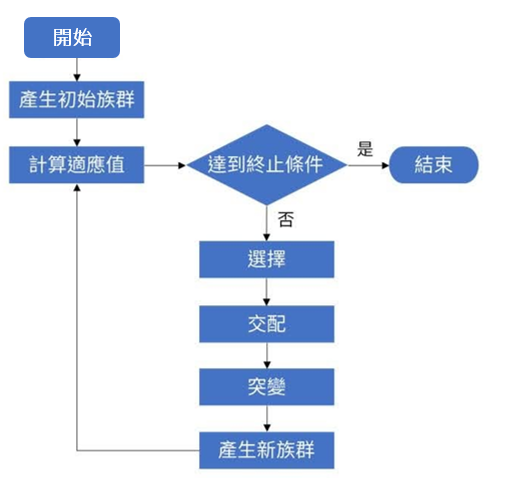


圖3-9基因演算法流程圖

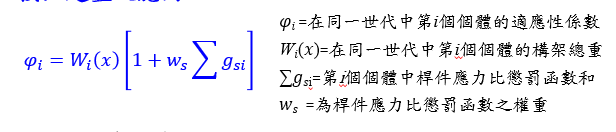


圖3-10基因演算適應性函數說明

3-4鋼構最優化設計-結果呈現

本章節我們以營建設計處提供的八層樓梯鋼構作為測試對象，進行基因演算法的優化。在測試過程中，我們注意到突變率對演算法的影響，當突變率小於10%時，演算法的收斂速度較慢；而當突變率大於等於10%時，演算法的收斂速度明顯加快。在突變率介於20%至50%的區間，演算法的收斂速率保持穩定。

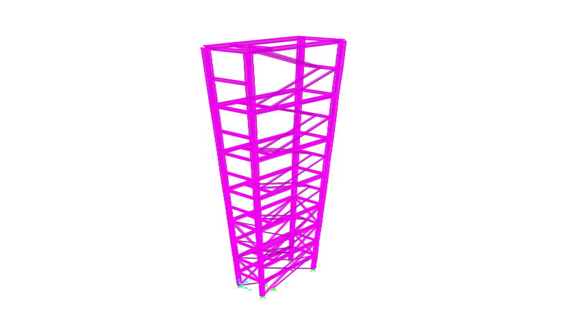
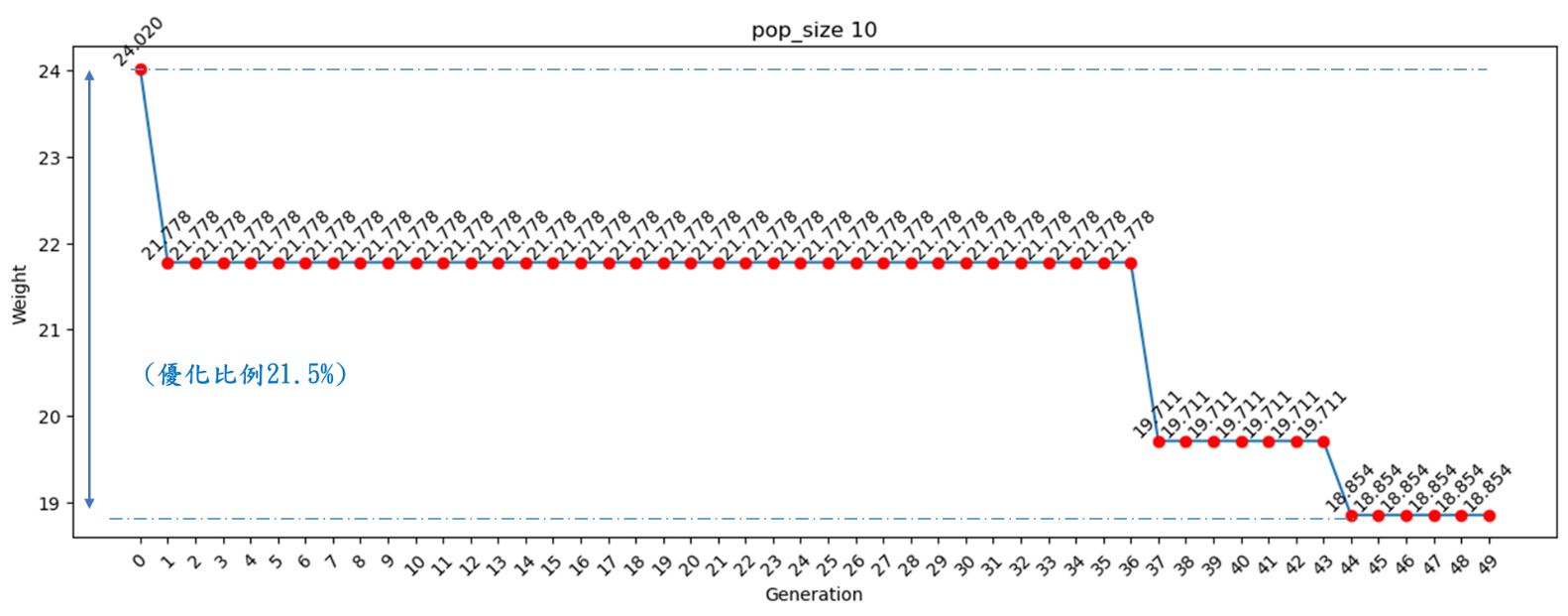
基於以上觀察，後續測試將突變率保持在20%的固定值，並針對其他參數，如族群數量、演化代數進行進一步測試，以探索最佳組合。透過系統性的測試，期望能找到最適合的參數設定，以提高基因演算法的收斂效率和優化結果。

圖3-11八層樓梯鋼構示意圖

3-4-1鋼構最優化設計-族群數量比較

首先，我們將繁衍世代固定為50世代，並將族群數量分別設置為10、50、100和500，以進行比較評估。根據圖3-12至圖3-15的結果顯示，當族群數量小於50時，收斂效果相對較差，僅有21.5%的降幅。然而，當族群數量為50、100和500時，其區間收斂效果分別為29.3%、29.5%和29.5%。可見，當族群數量達到一定基準值後，對於鋼構重量的收斂效果相近，族群數量50和族群數量100、500之間僅有0.2%的差異。主要的差異在於收斂速度，族群數量為50時，在演化至第20代時鋼構重量降至16.974，族群數量為100時，在第18代時重量降至16.962，族群數量為500時，在第17代時重量降至16.964。由此可見，當族群數量較大時，收斂速度較快，但當達到一定數量時(測試結果為100~500區間)，族群數量對於收斂速度的影響逐漸下降。

值得注意的是，在運行此模型時，當族群數量為50時，每一世代的計算時間約為8分鐘，族群數量為100時，每一世代計算時間為15分鐘，可以看出當族群數量增加時，計算時間也隨之增長（族群數量為500時，每一世代計算時間約為1小時）。基於綜合考量，為了方便後續測試，我們決定將族群數量固定在100，以保持優異的收斂效果同時大量減少後續測試所需時間。

圖3-12 基因演算法-族群數量10

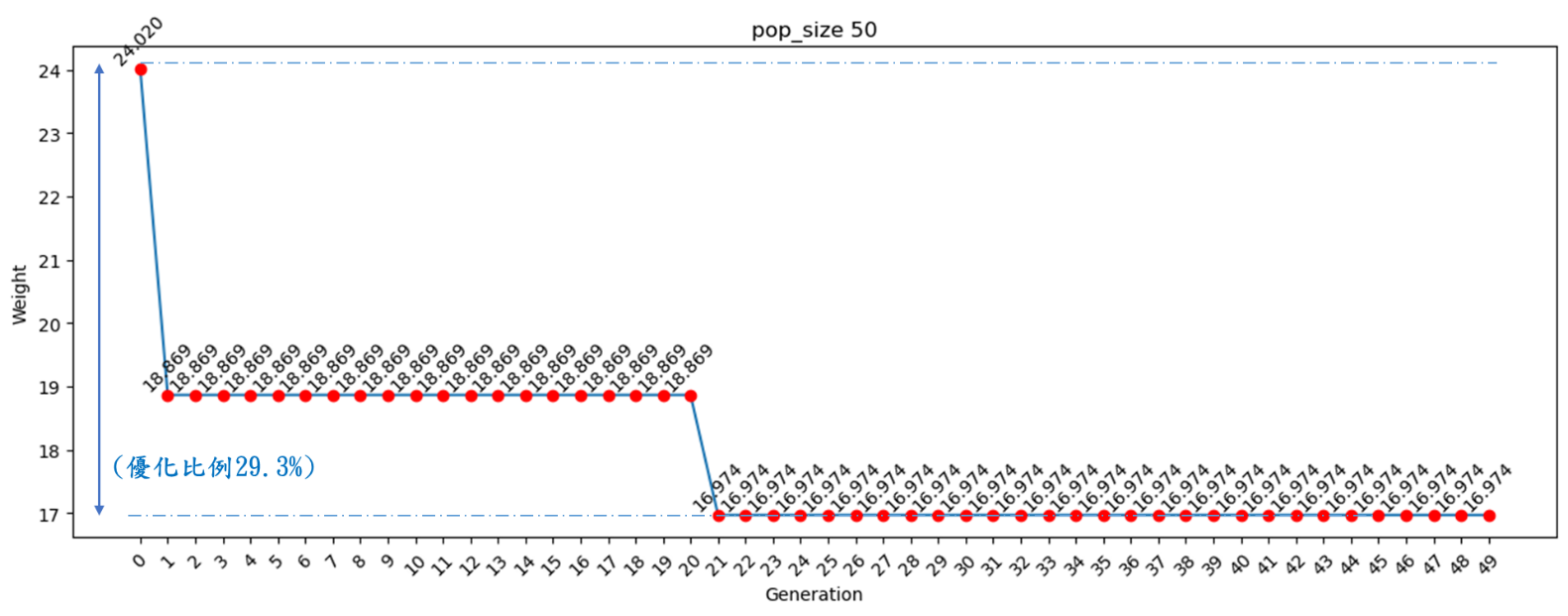
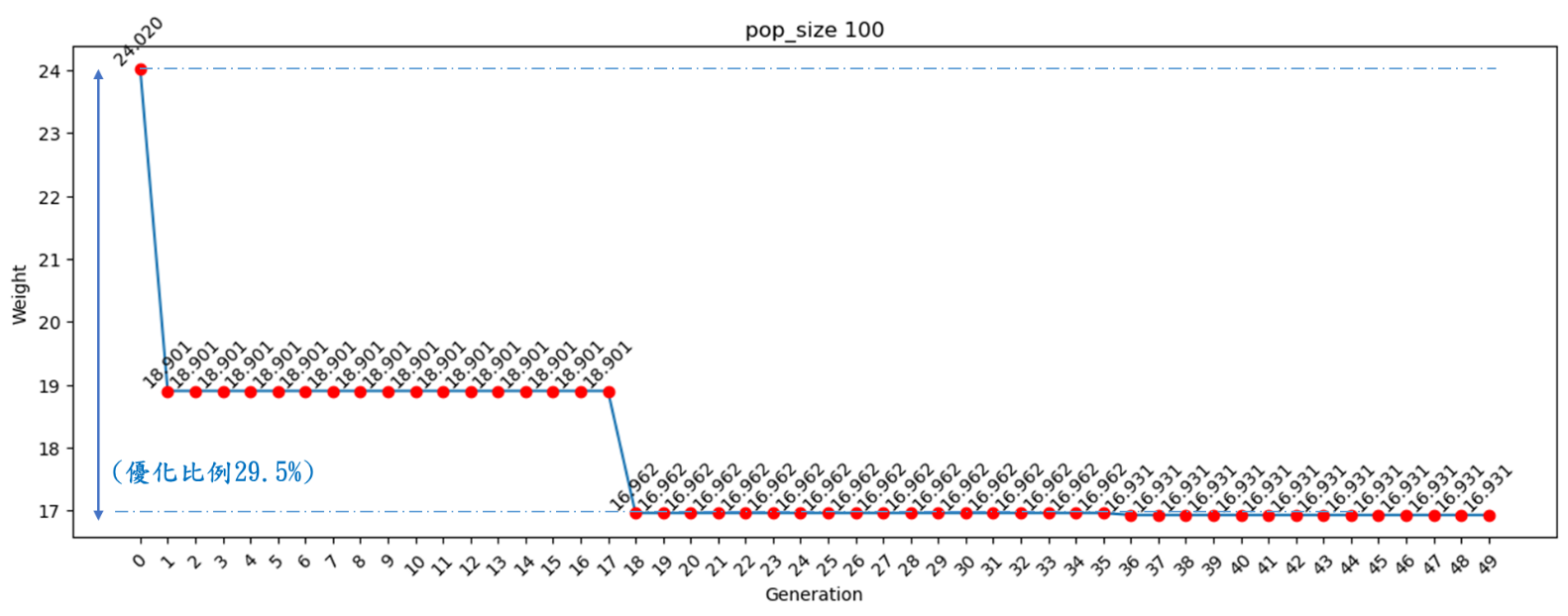


圖3-13基因演算法-族群數量50

圖3-14基因演算法-族群數量100

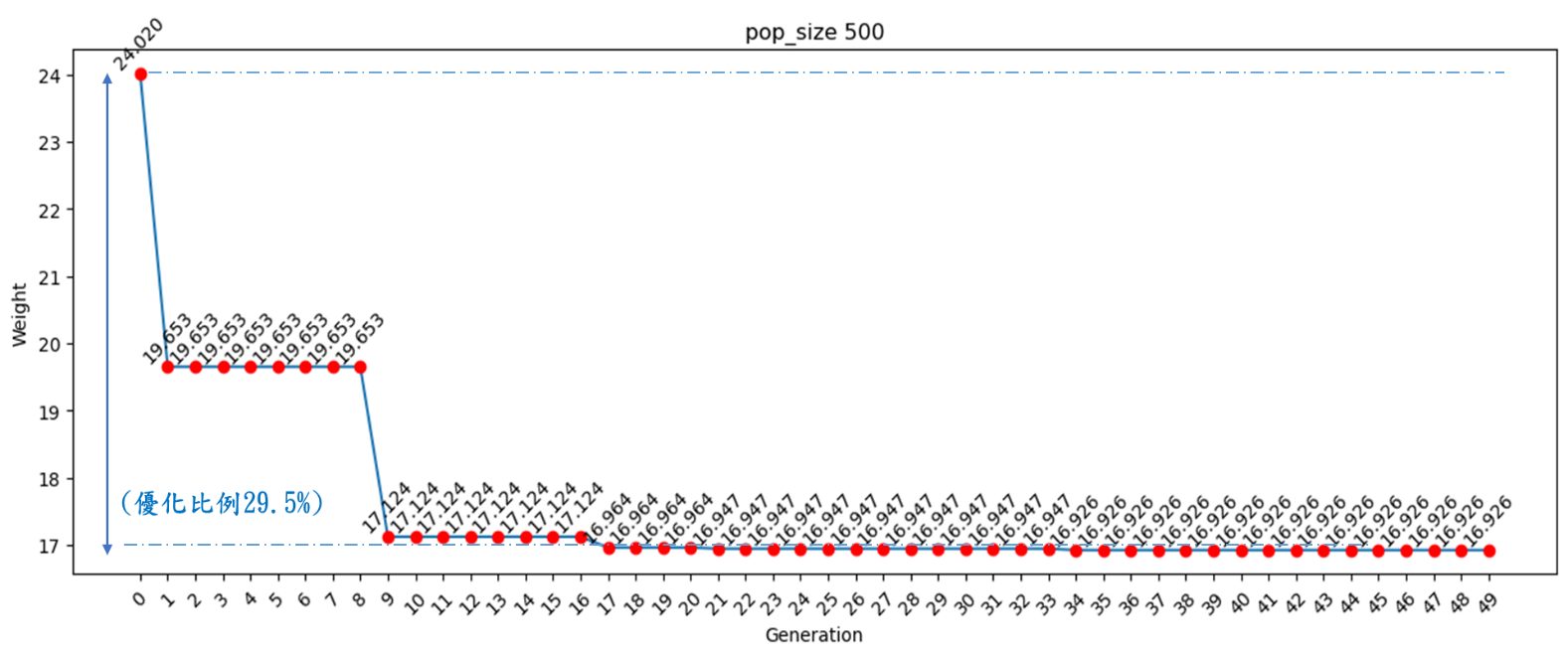


圖3-15基因演算法-族群數量500

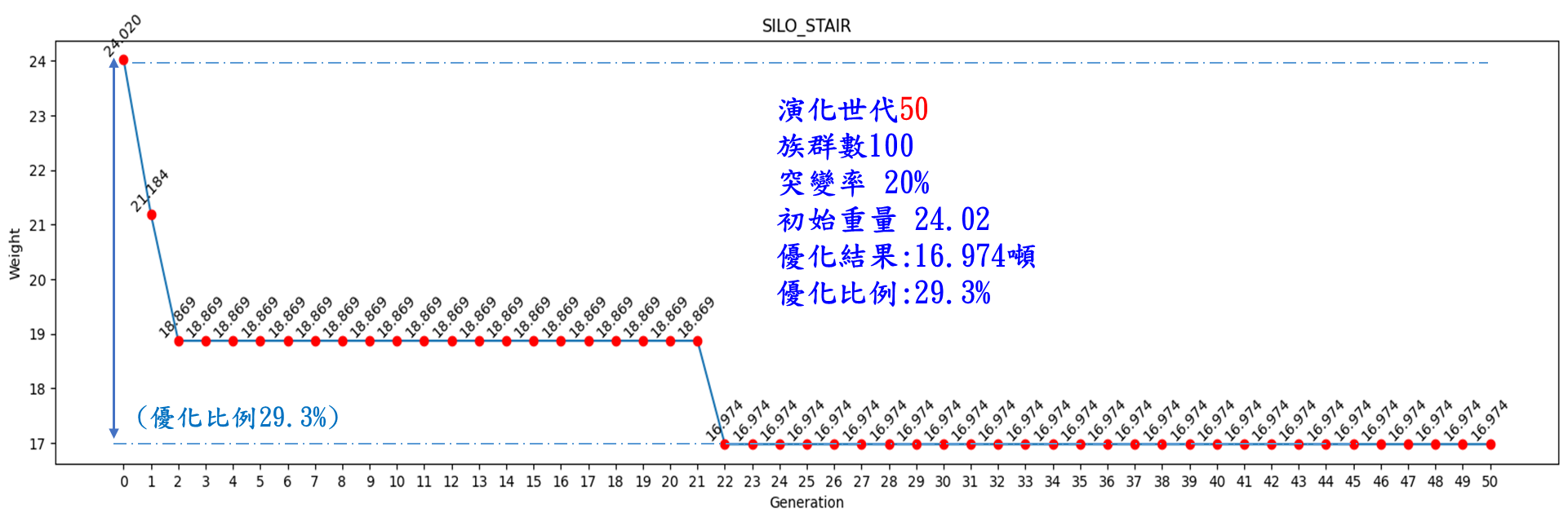
**3-4**-2**鋼構最優化設計-演化代數比較**

在3-4-1小節講族群數量定在50個個體數後，本小節我們著重探討族群數量對最佳化結果的影響，並進行了不同演化世代的比較。比較演化世代為50、100、200和500的分別進行實現，並保持族群數在每次實驗中都為100個個體，突變率則固定在0.2%，其結果如下圖3-15~圖3-19所示。

首先，在演化世代為50的結果中，我們觀察到優化結果達到了16.974噸，相對於初始中量的改進達到了29.3%，接著，我們將族群數保持在100個的設定下進行了演化世代為100、200和500的實驗。在這些實驗中，優化結果分別為16.974噸、16.992噸和16.983噸，同樣保持在約29.3%的改進幅度。可看出後續演化世代增加並未帶來明顯改善，顯示出在演化世代達到50代時，優化效果已達到相對穩定的水平。

這些結果表明，在案場中，隨著演化世代數的增加，最終優化結果重量有約29.3%的差異，且在演化世代為50時優化效果已趨於穩定。

在與營建專業人員討論後，我們確定結構已達到最佳化且符合現場施工方便性之設計需求。也代表設計方案不僅在理論層面上可行，同時也能適應實際施工環境的需求。

圖3-16基因演算法-演化世代50

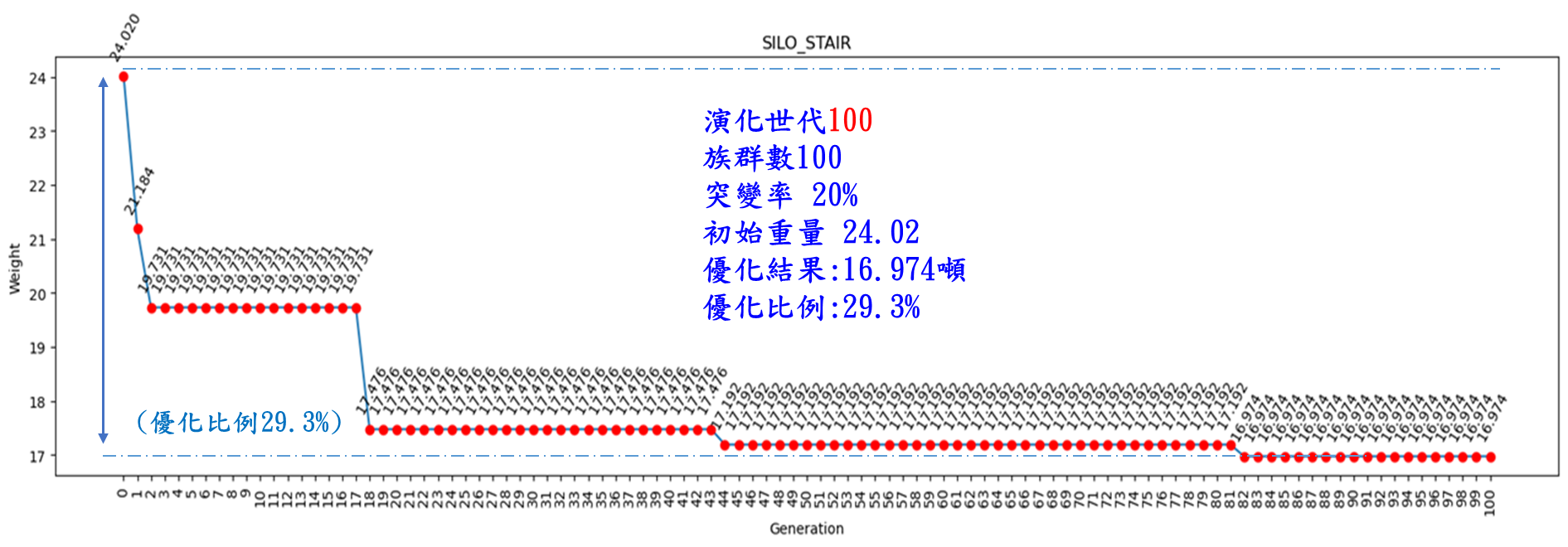
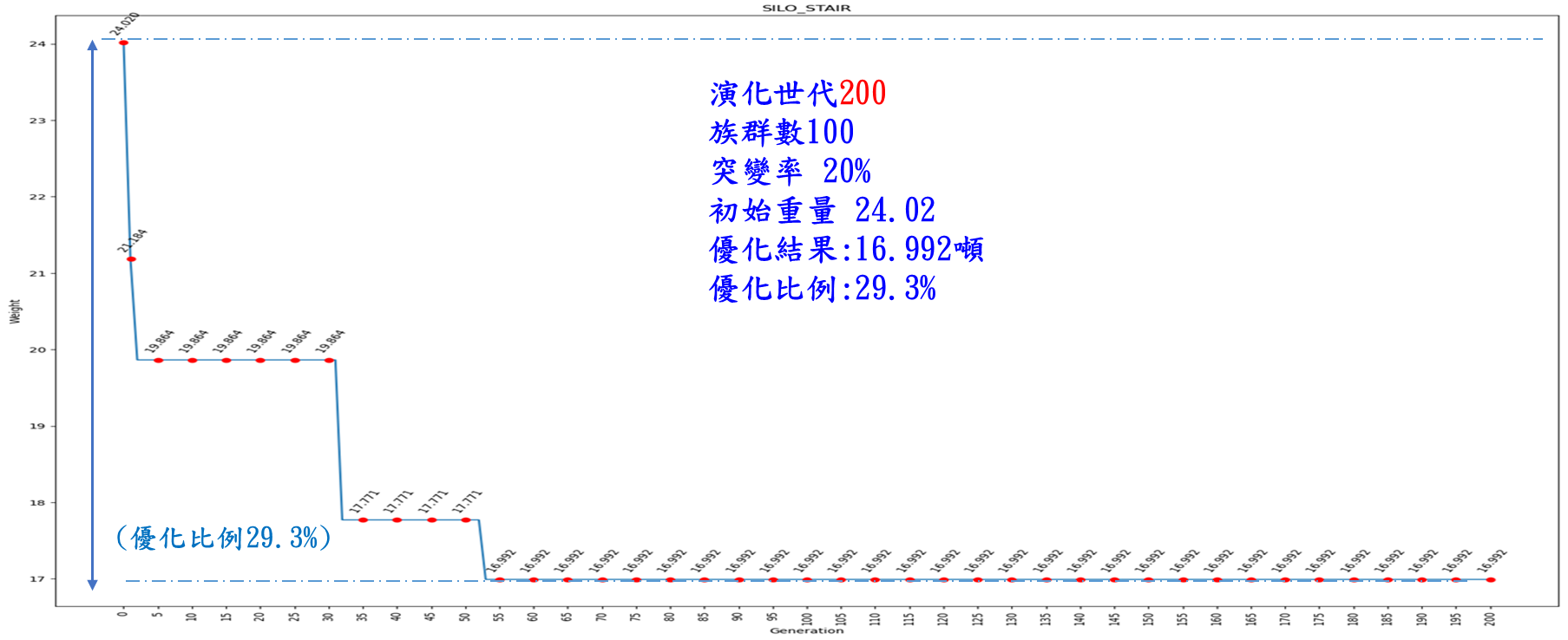


圖3-17基因演算法-演化世代100

圖3-18基因演算法-演化世代200

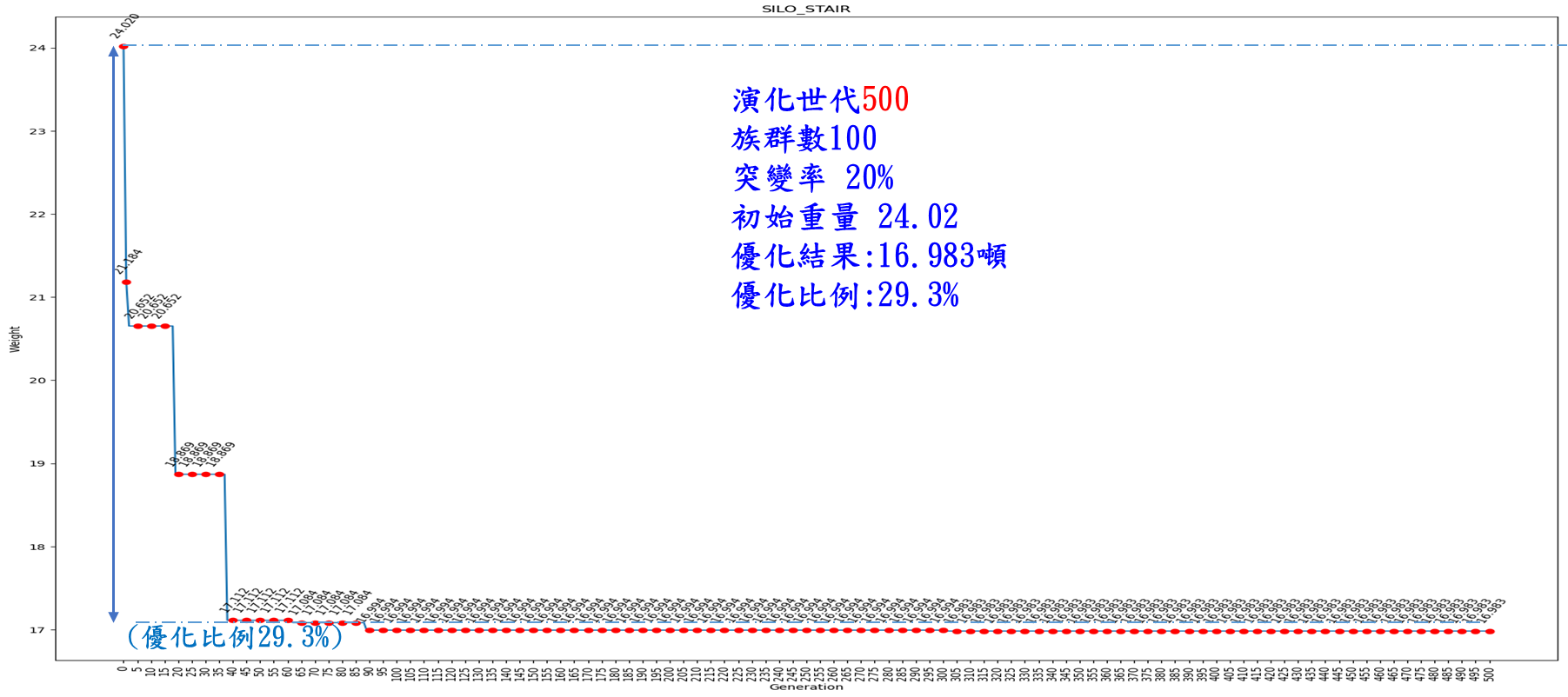


圖3-19基因演算法-演化世代500

1. 工作心得

轉眼間，來到自控處已經過一年，在近三個月間期間我承擔了「ARO2廠甲苯塔優化模組人機介面建置」和「鋼構最優化設計AI模組開發」兩項工作，讓我深刻體會到在工程領域中應用AI技術的重要性和挑戰。

第一項案件中，建置可視化工具，以協助單元製程的操作調整並提供節能操作條件建議，以確保最佳能耗操作並符合品質管制基準。在與營建專業人員密切合作的過程中，我們發現了甲苯塔C670操作的保守性，進而發現了節能優化的潛力。我們開發了AI模組，架設了Dashboard，以互動式網頁的方式提供現場人員操作判斷依據。這項工作讓我深刻感受到AI在工程實踐中的巨大優勢，不僅能提高操作效率，同時為節能環保做出積極貢獻。

第二項案件則聚焦於鋼構設計的最優化。我們結合機器學習和基因演算法，開發自動化的優化設計程式，取代了傳統的人工試誤法，大大節省了時間和人力成本。這項工作讓我深入了解了AI演算法的運作原理，並在實際應用中體驗到了其優越性。透過這個自動化過程，我們能夠找出最佳的結構設計，實現了經濟節能和減少碳排的目標。

總結而言，此次案件開發不僅著重於數據和理論模型，同樣也強調與專業實務的結合，讓其能上線應用。這種跨領域合作與交流有助於確保我們的研究成果能夠真正對於實際工程和項目產生積極影響。

在這些工作中，我不僅學習到了專業的知識和技能，同時也體驗到了與專業人員密切合作的重要性。他們的專業見解和實踐經驗對我們的研究和開發起到了關鍵的指導作用。此外，這些工作也讓我進一步了解到AI技術在工程領域的無限潛力，並鞏固了我對未來工作的熱情和動力。