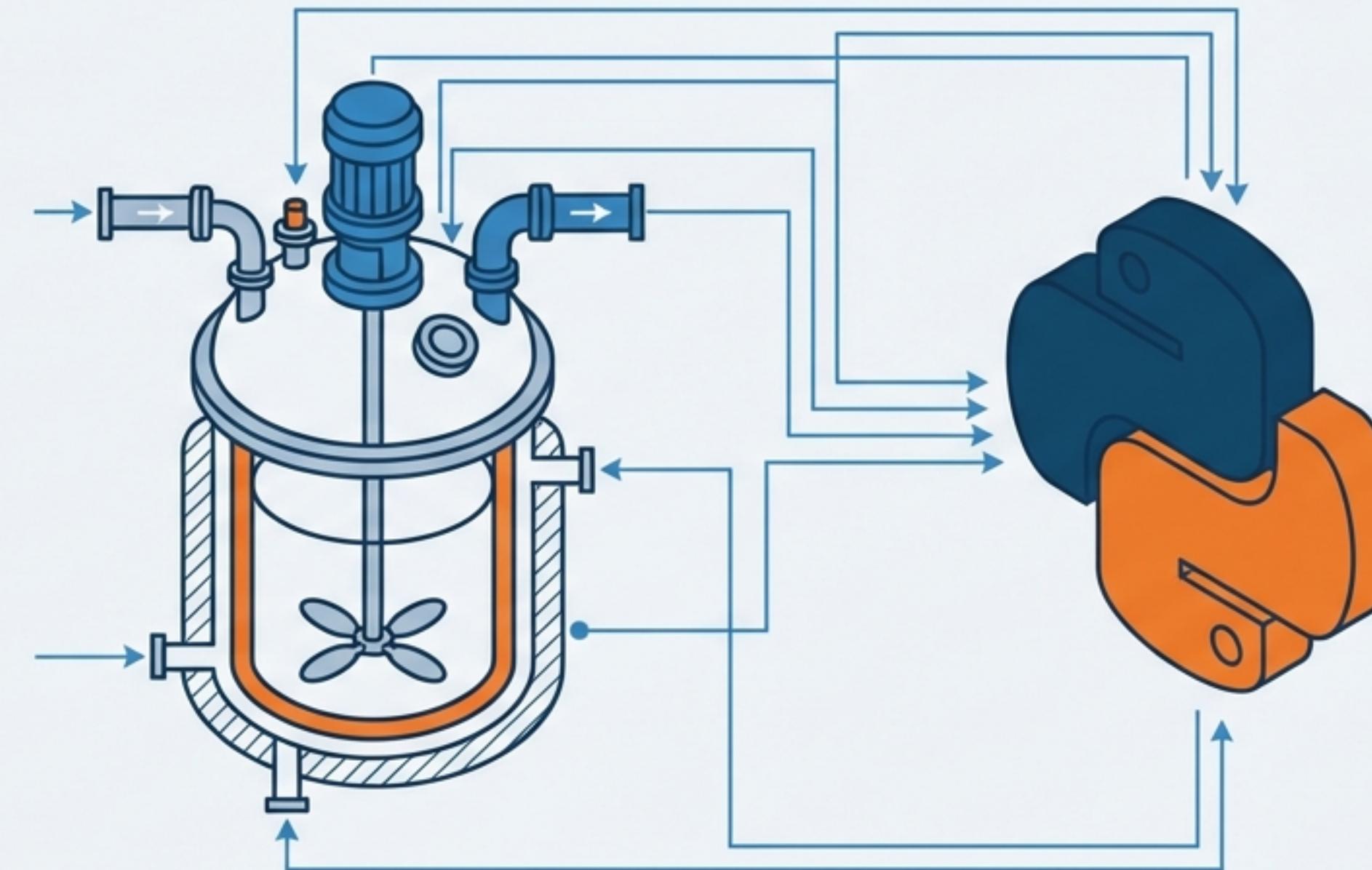


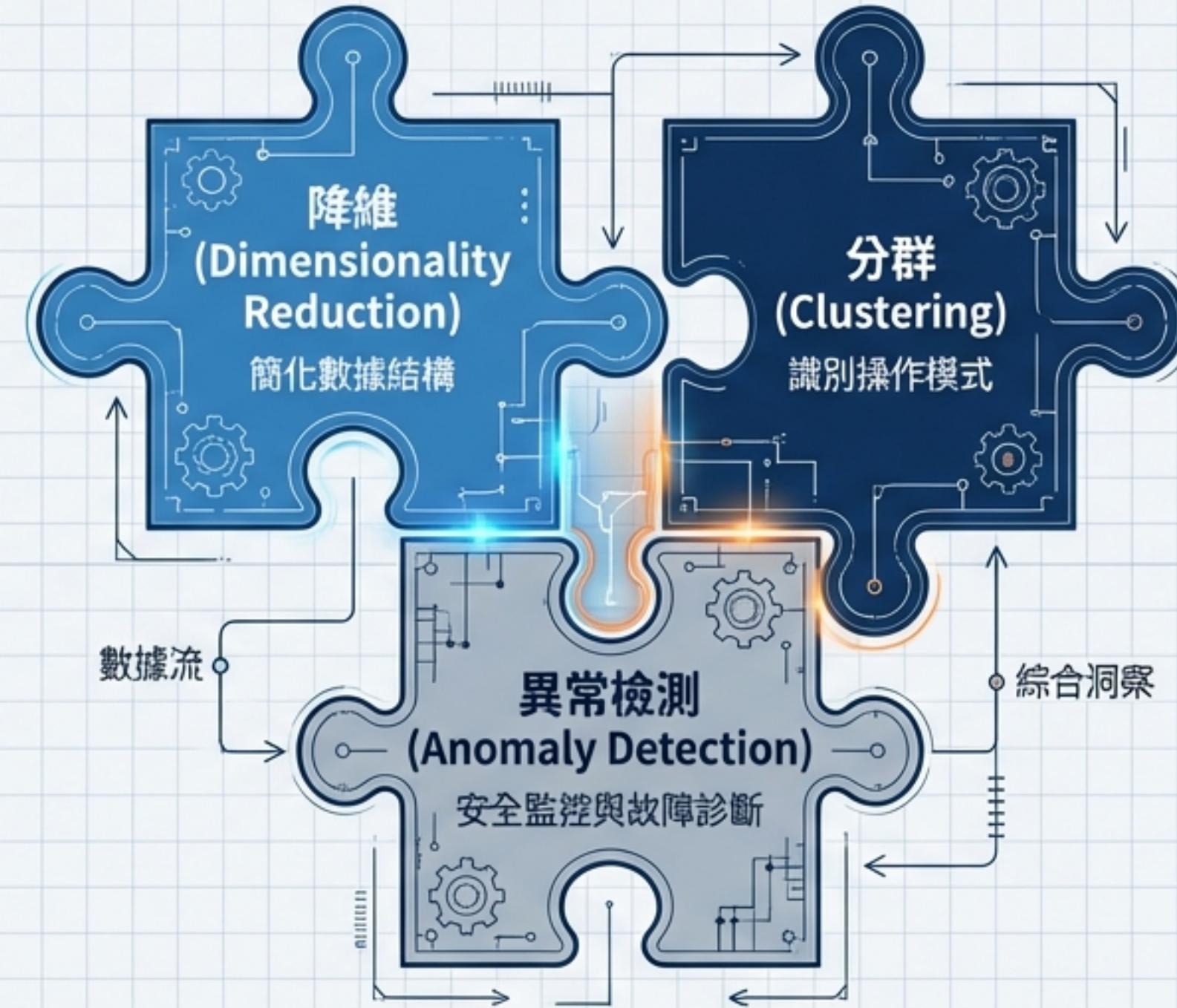
Unit 09 綜合案例研究與方法組合實務應用

以田納西-伊士曼製程 (TEP) 為例：從數據降維到故障診斷的端對端流程



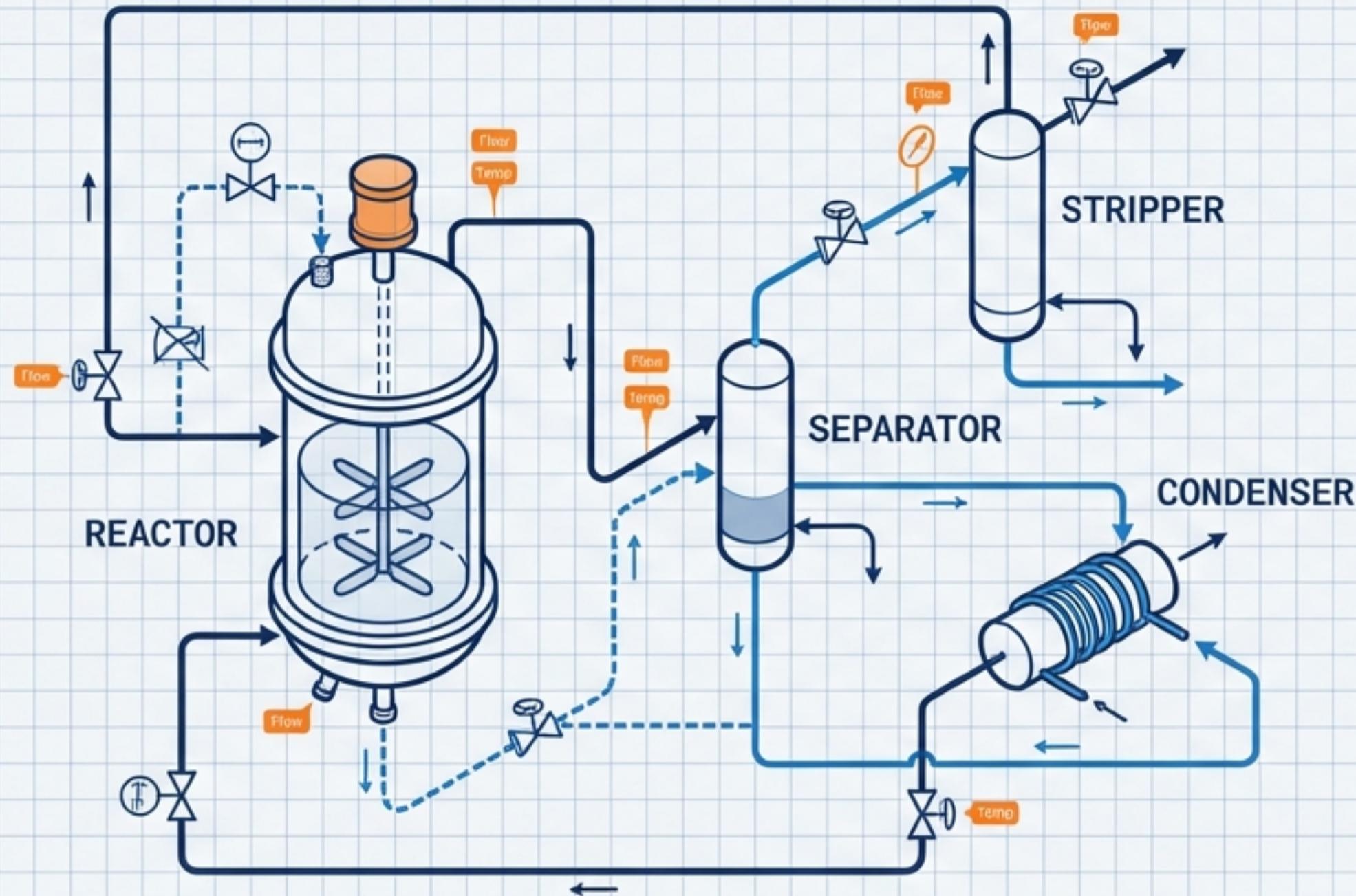
為什麼需要綜合分析？單一方法的侷限與整合優勢

-  **互補性 (Complementarity)**：結合不同視角，獲得全面洞察（如：降維後更有利於分群）。
-  **驗證性 (Validation)**：不同方法的結果可相互參照（如：異常點是否位於特定群組邊緣）。
-  **穩健性 (Robustness)**：透過多重檢測降低誤報率，提高工業應用的可靠度。



Insight：在真實化工製程中，變數間存在強耦合與非線性關係，唯有「組合拳」才能有效解決複雜問題。

案例對象：田納西-伊士曼製程 (TEP)



Data Specification

基準地位

Downs & Vogel (1993) 提出，製程控制與故障診斷的標準 Benchmark。

反應機制

$A(g) + C(g) + D(g) \rightarrow G(\text{liq})$
Product G & H 為主產物

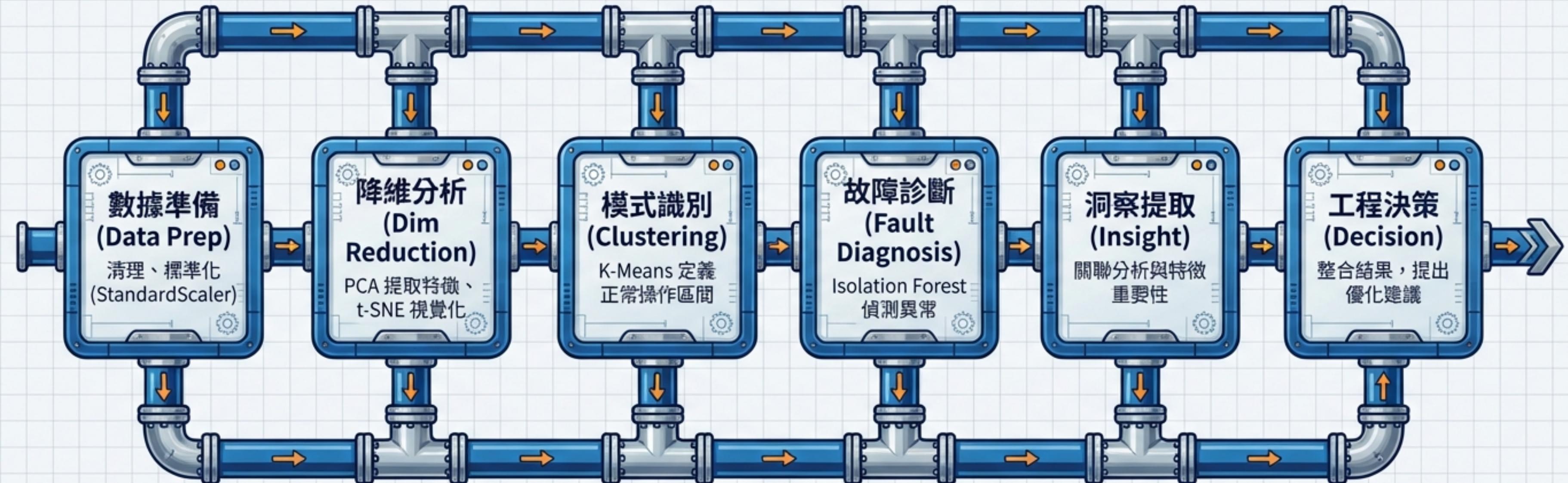
數據規模

52 個變數 (41 個測量變數 XMEAS +
11 個操作變數 XMV)

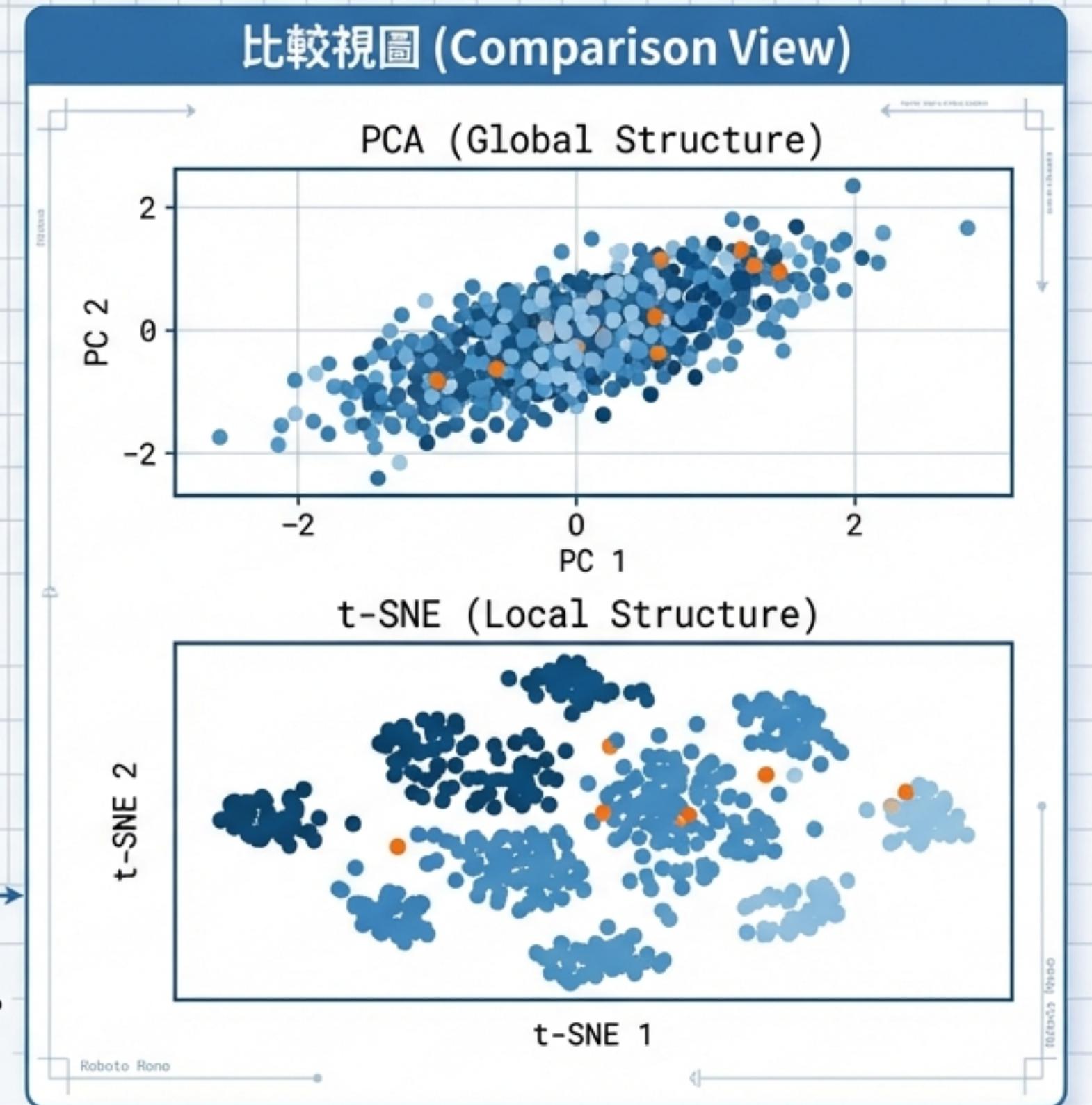
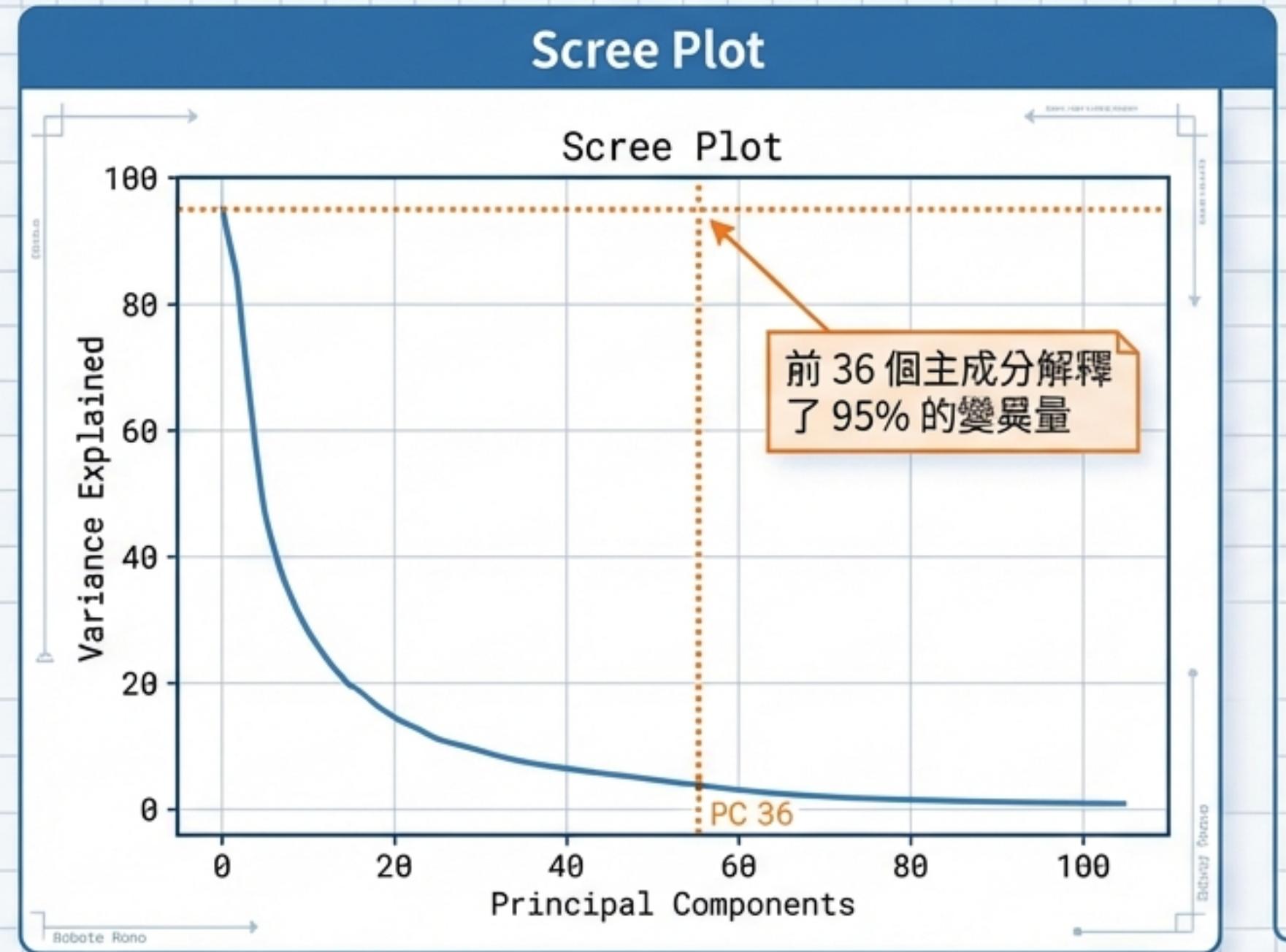
挑戰場景

包含 21 種預定義故障 (IDV 1-21) 及
強耦合、非線性動態特性

數位工程師的藍圖：綜合分析工作流程

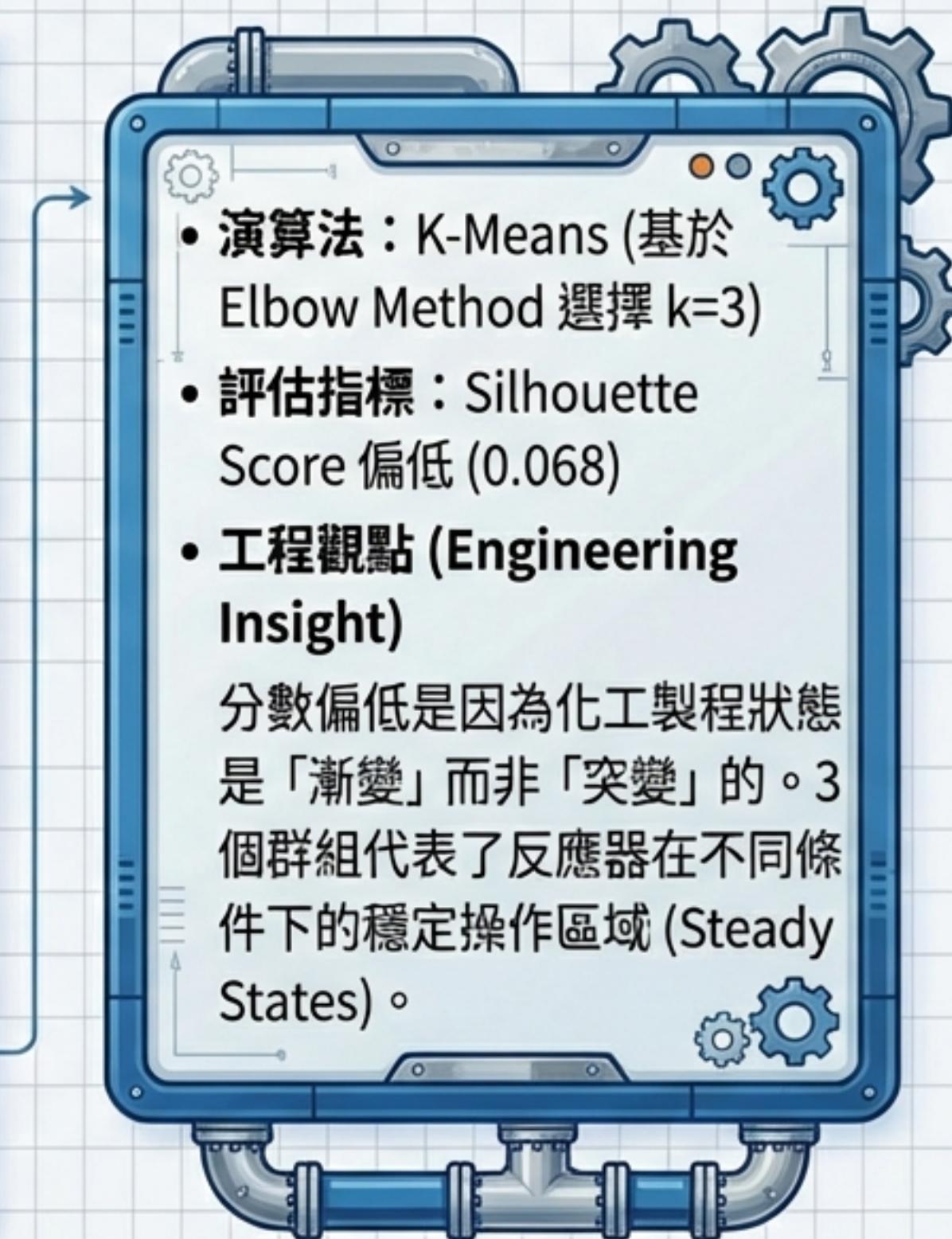
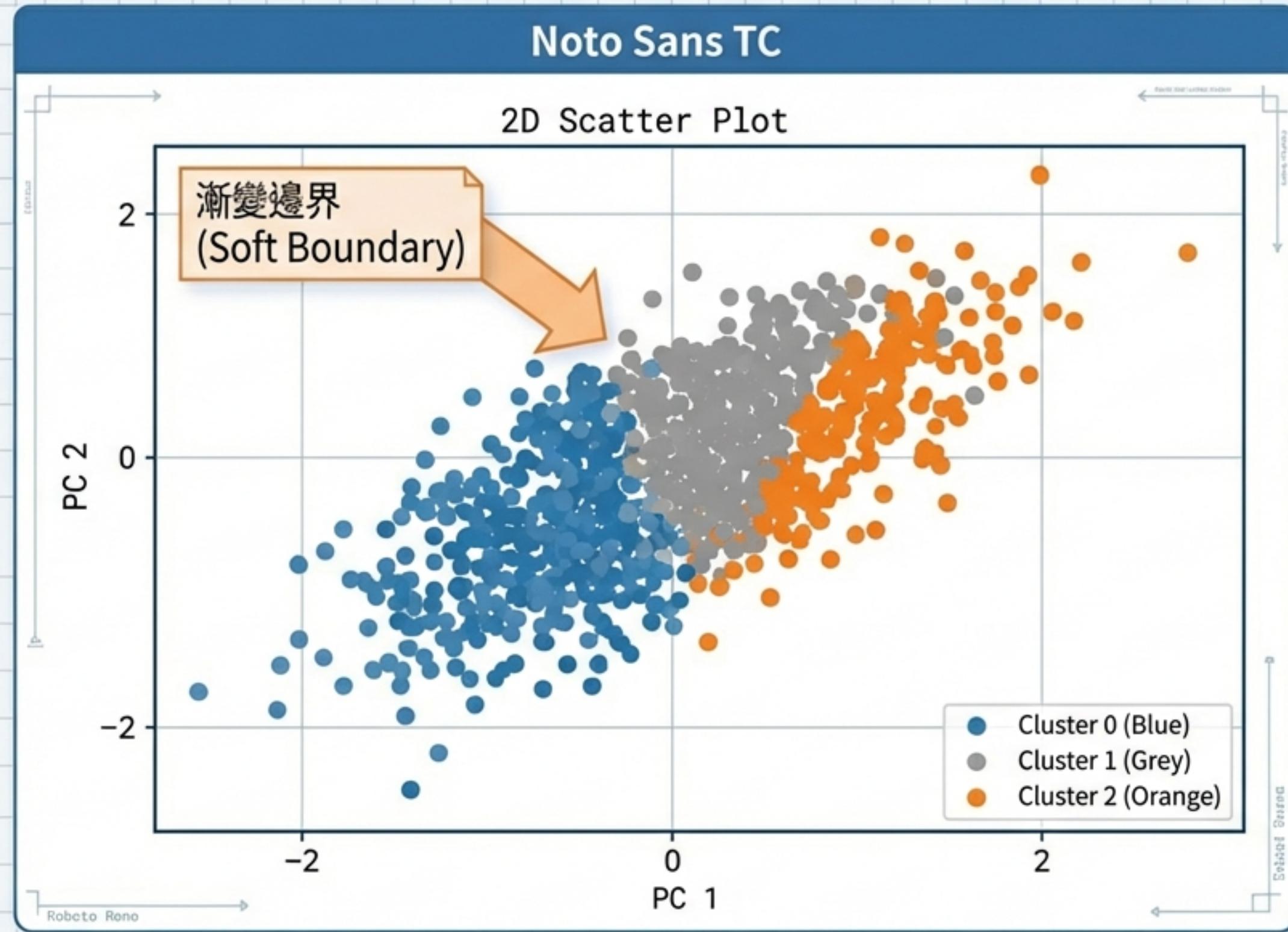


駕馭複雜度：降維與特徵提取



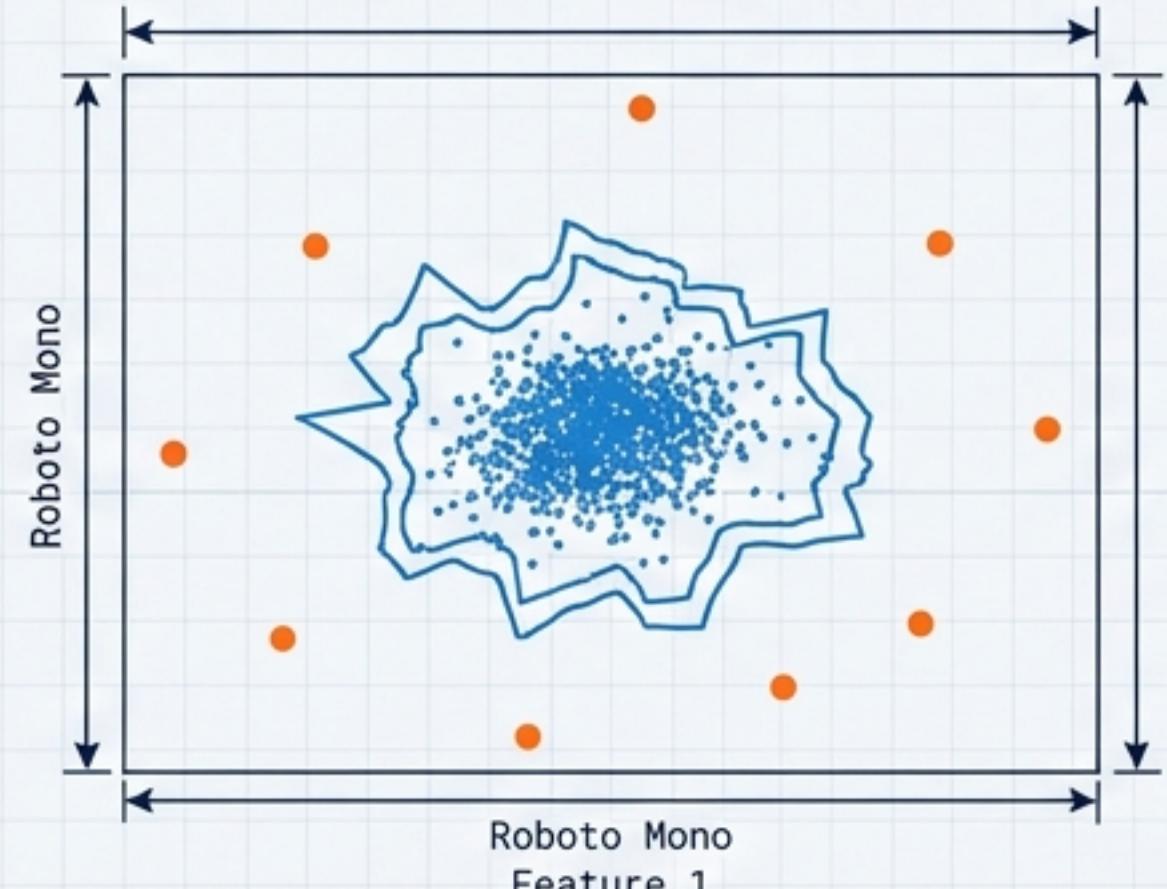
- PCA 優勢：保留全局結構，適合特徵提取與監控 (Process Monitoring)。
- t-SNE/UMAP 優勢：擅長捕捉局部結構，適合視覺化探索數據群聚。

識別操作模式：分群分析 (Clustering)



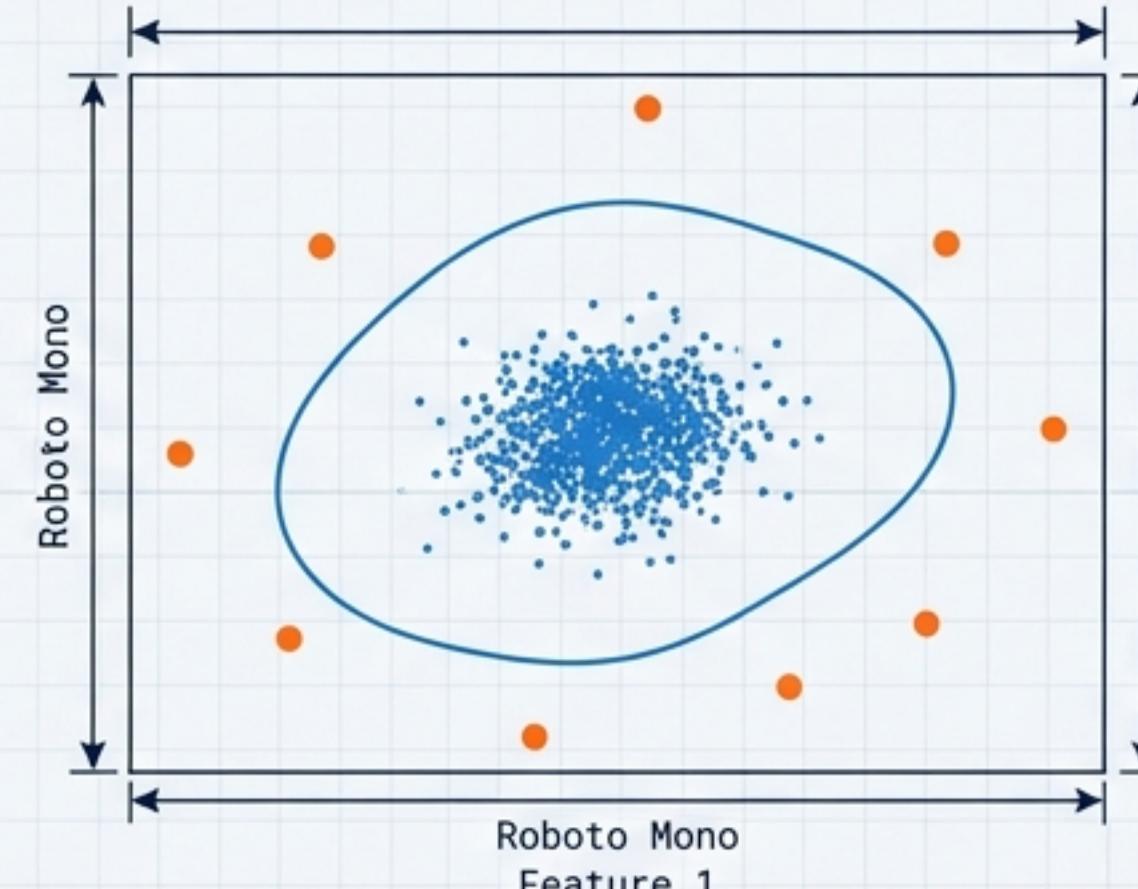
數位安全連鎖系統：異常檢測

Isolation Forest (孤立森林)



利用隨機切割樹，異常點路徑短。適合高維數據。

One-Class SVM



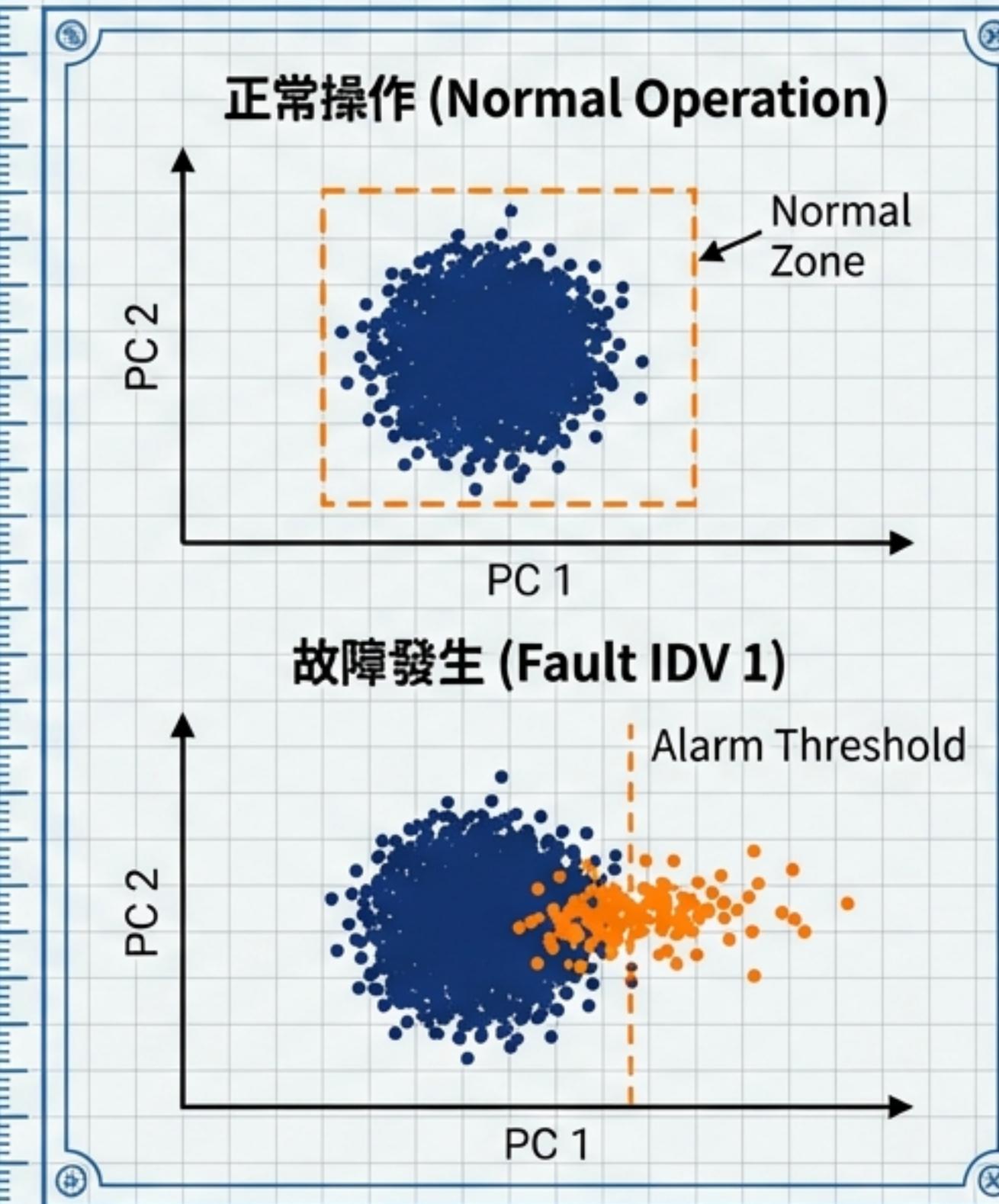
尋找包圍正常數據的超平面。定義嚴格邊界。

Key Stats Box

設定閾值 (Contamination) : 0.05 (5%)
準確率 (Accuracy) : ~82% (Isolation Forest)

實戰演練：診斷 IDV(1) 故障場景

故障場景：IDV(1) - A/C
進料比例改變 (A/C Feed
Ratio Change)。

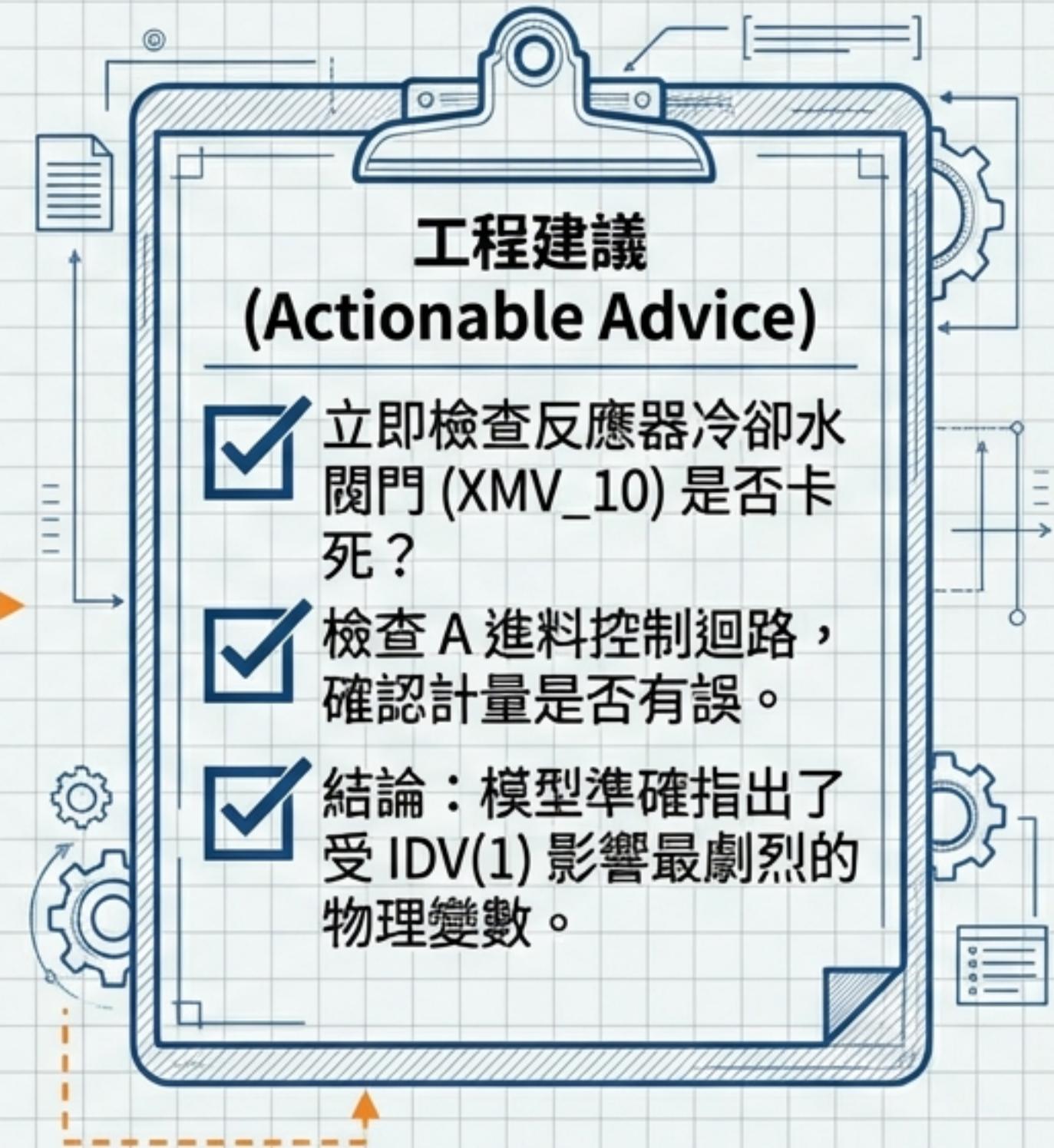
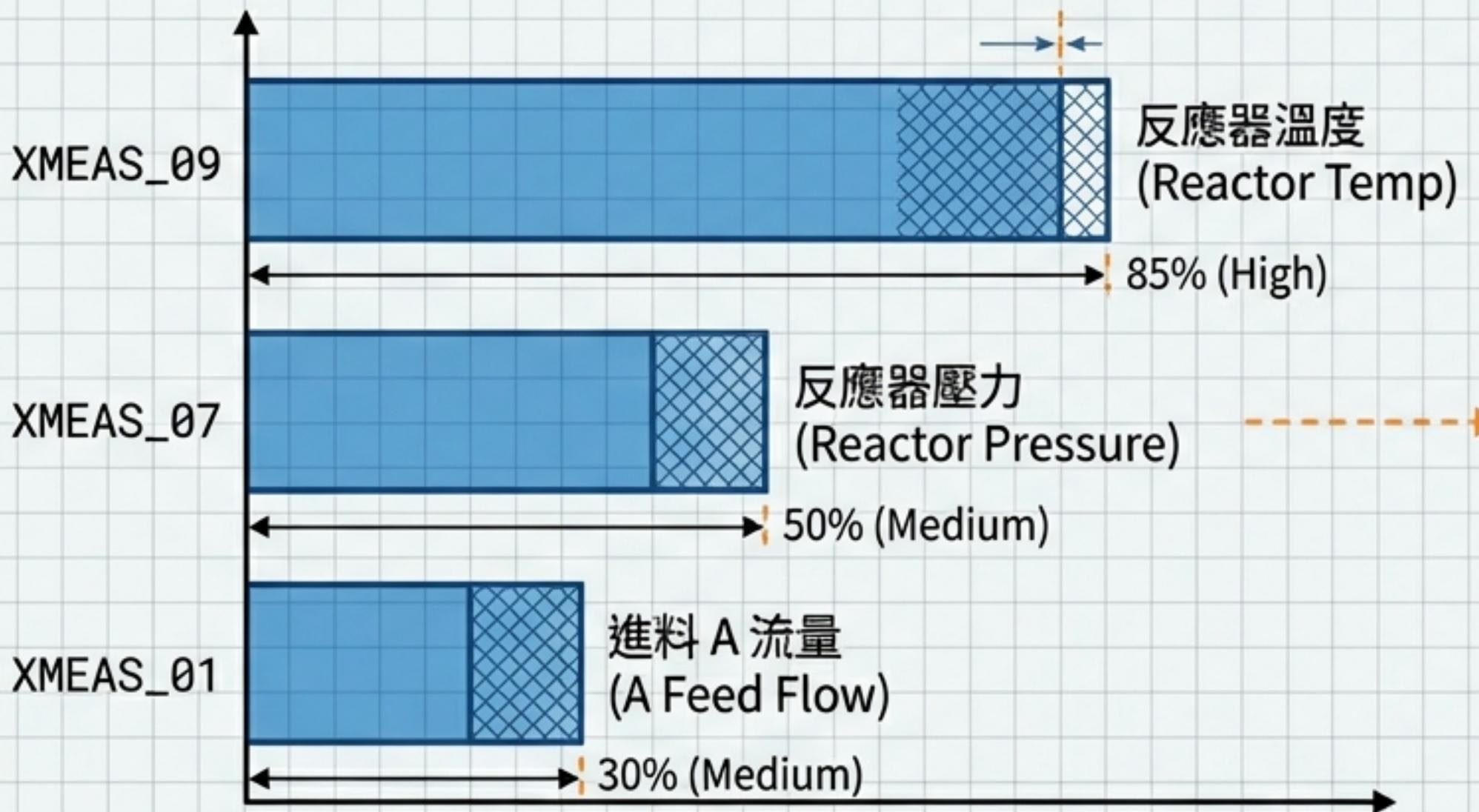


診斷結果

- * 模型成功偵測到數據漂移。
- * **Elliptic Envelope** 準確率最高 (85.1%)
- * **Isolation Forest** 次之 (81.9%)
- * **Recall (召回率) > 0.85**，符合「寧可誤報不可漏報」的安全原則。

從數據到決策：特徵重要性與根因分析

Top Contributing Variables



工程師的工具箱：方法選擇指南

製程監控 (Monitoring)



首選: **PCA**

原因: 保留全局結構，易於計算
 T^2/SPE 統計量。

故障檢測 (Fault Detection)



首選: **Isolation Forest**

原因: 處理高維數據快，對異常敏
感。

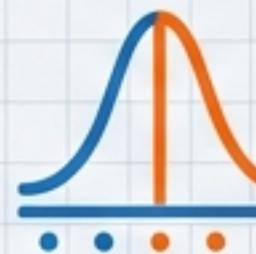
操作模式 (Op. Modes)



首選: **K-Means / GMM**

原因: K-Means 適合已知群數，
GMM 適合邊界模糊。

數據分布已知



首選: **Elliptic Envelope**

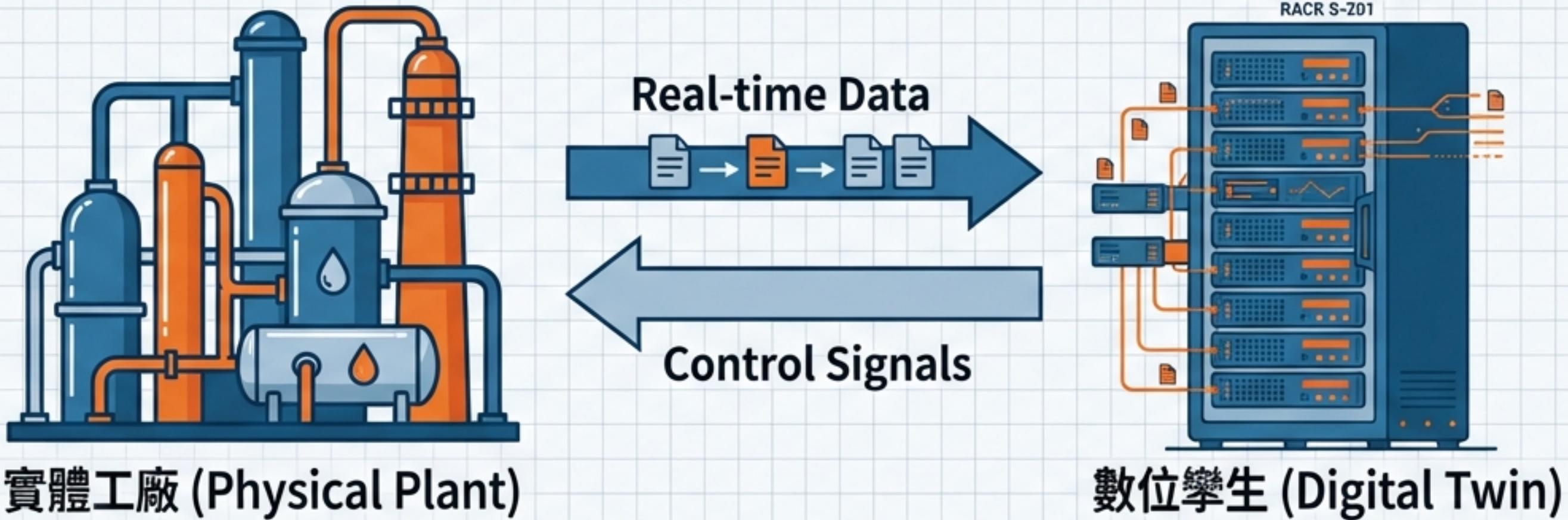
原因: 若數據符合高斯分布，效果
最佳。

Rule of Thumb: 先用 **PCA** 理解結構，再用 **Isolation Forest** 建立防線。

實務挑戰與最佳實踐



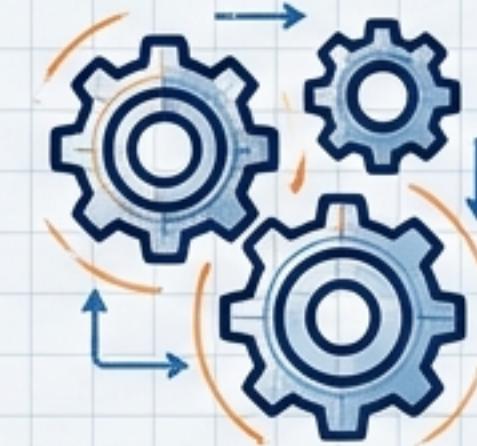
展望未來：從數據分析到智慧製造



1. 數位孿生 (Digital Twin)：利用 PCA/分群結果建立即時狀態映射。
2. 預測性維護 (Predictive Maintenance)：在 IDV(1) 導致停機前，捕捉微小信號漂移。
3. 自適應控制 (Adaptive Control)：根據 K-Means 識別的當前群組，自動切換 PID 參數。

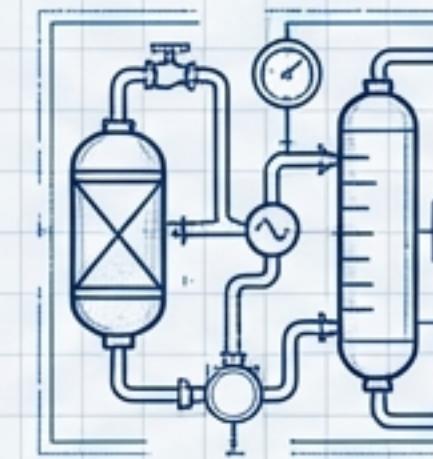
結語與下一步

整合是必須的



單一方法無法應對 TEP 等級的複雜度，必須打組合拳。

TEP 是標準



掌握 TEP 分析流程，即掌握了化工大數據的基礎。

領域知識是核心



數據科學工具必須服務於工程目標 (Safety & Efficiency)。

→ 下一步 (Next Step):

開啟並執行 Jupyter Notebook: `Unit09_Integrated_Case_Study_Overview.ipynb`
親自動手調整 `contamination` 參數，觀察對故障檢測率的影響。