

具整體學習機制之在製品品質檢測暨缺陷分類技術

WIP Quality Detection and Defect Classification based on Ensemble Learning Methods

鄭光宏

夏啟峻

葉峻賓

江宜霖

林順傑

蔡煥文

Kuang Hung Cheng, Chi-Chun Hsia, Jun-Bin Yeh, Yi-Lin Chiang, Shun-Chieh Lin, Huan-Wen Tsai

中文摘要

產品品質是企業追求客戶滿意、提高競爭力的重要指標，生產檢測則是確保產品品質的重要手段，然為節省成本及時間，目前業界採取抽檢方式來進行品質監測，但如此並無法確保每一個產品品質，故透過資料分析手法來改善工廠之生產品質與效率是一重要課題。本研究所提出技術採用機器學習方式進行大數據資料建模，包含支援向量機、邏輯斯迴歸及決策樹等各式分類器，在普遍製造現場資料特性未知情形下，以袋式(Bagging)整體學習(Ensemble Learning)機制進行模型訓練，有效降低傳統方法因選擇錯誤分類器而造成過適(overfitting)風險，達成穩健的品質分類效能，以低時間延遲方式，檢測產品的品質與良率，當不良品數目過高或良率不佳時，警示現場作業人員進行必要的停機檢查。導入本技術可幫助現場人員即時掌握在製品品質狀況，防杜廢品進入下一到製程而產生之二次損失成本，同時確保產品品質。

Abstract

In order to gain customer satisfaction and enhance competitiveness, manufacturers usually aim to offer their clients with high quality products. However, under the consideration of time and cost reduction, random sampling has become a popular method for production quality metrology whereas not all the products will be examined and assured. To overcome this problem, this paper proposes a WIP Quality Detection and Classification method based on Ensemble Learning with well-known machine learning approaches, like Support Vector Machine (SVM), Logistic Regression, Decision Tree, and so on. Unlike a statistical ensemble in statistical mechanics, which is usually infinite, a machine learning ensemble refers only to a concrete finite set of alternative models, but typically allows for much more flexible structure to exist among those alternatives. Therefore, while most of the manufacturing data characteristics are unknown (means infinite hypothesis), our method can reduce traditional model bias from overfitting to enable real-time product quality detection with achieving stable classification efficiency. Once the yield rate is below certain level, the system will send a warning notice to inform the process engineer to shut down the process and cut the costs of rework, such as correcting of defective, failed, or non-conforming item, during or after inspection.

關鍵詞(Key Words)

整體學習(Ensemble Learning)

監督學習(Supervised Learning)

品質分類(Quality Classification)

在製品品質檢測(WIP Quality Detection)

1 · 前言

因應全球社會與經濟環境的變遷，各國均注意到製造業對於就業與經濟成長的重要性，分別依據國家所擁有之領先優勢，對未來自我之製造業競爭力提出因應政策。基於資通訊技術的進步，運用資通訊技術使得產品於設計、生產之過程更有效能的智慧製造成為主流趨勢，各國政府競相投入資源改造現有之製造環境，例如德國在2012年提出「工業4.0」政策，導入以整合資通訊軟、硬體，並結合物聯網概念的虛實化生產系統 (Cyber-physical System, CPS) 來推動其智慧製造優化基礎；美國則是啟動AMP(先進製造合作夥伴計畫, Advance Manufacturing Partnership)計畫，發展數位製造、先進製程與3D列印等項目來提升競爭力，日本、中國亦發表其製造業改造政策，全球因此激起了「再工業化」的浪潮。而「工業4.0」所運用之資通訊技術由物聯網、巨量資料、雲端服務等三項技術組成，故除了透過物聯網蒐集現場/製造資料，並將資料匯整於雲端平台外，透過資料分析手法來改善工廠之生產品質與效率將是一重要課題 [1]。

在成大製造所鄭芳田講座教授於科學發展期刊的「全自動虛擬量測」 [2] 文中亦提到，國際半導體技術發展藍圖中把虛擬量測這種資料分析技術列為未來工廠資訊與控制系統內的一項重要技術。2009 年在美國舉行的先進機台控制與先進製程控制國際研討會中，IBM 公司在題目為「Solar APC」的演講中也表示，虛擬量測可協助太陽能產業提升製造過程的良率。台積電在2013 年台灣歐盟資通訊科技合作策略會議內有關下世代 (即18 英寸或450 毫米) 晶圓的議題所發表的「450毫米晶圓轉型：挑戰與期望 (The 450mm Transition: Challenges and Expectations)」演講中表示，可藉虛擬量測技術使下世代450 毫米生產機台具有現場自我監測、自我控制和減少監控晶圓消耗等的能力，係故如虛擬量測之品質分類技術可在如半導體、面板或太陽能產業的產品尚未或無法進行實際量測的情況下，利用生產機台參數分類其所生產產品的品質，以進行線上且即時的產品品質分類 (以達到生產片全檢的效果)、機台

效能監控及生產製程改善，如此可即時發現異常，避免重大損失。

在其他產業亦可見到相同的需求，例如，鋼鐵業連鑄製程是鋼液從1600°C的液態，降溫至600°C以下，結晶成為鑄胚的過程，會有許多感測器量測到的製程資料，包含模液面高度、冷卻水的水量/溫差、特定氣體的流量、合金元素的成份比例...等，而這過程中隨時都會有品質變異機會。由於其鑄胚品質量測成本高(ex: 鋼鐵業耐高溫環境)及量測耗時(ex: 鋼鐵鑄胚量測需經冷卻，空冷要7天，水冷要3天)，故目前大多以人工抽檢方式進行，以致掌握在製品品質不易，讓不良品流入後段製程。另外，在研磨製程上往往因砂輪供應品質不穩定，在標準保養程序實施下，仍會發生砂輪研磨的效能變異，造成的工件報廢成本很高。業者需要能透過製程資料達到即時的分析並反饋資料給現場做立即的改善之機制，並將現行採用之管制與規格界限處理方式演進成更具效率之可隨加工環境調適的處理系統。

然巨量資料應用在全球各國發酵之際，伴隨而來的問題，就是各國普遍缺乏資料科學家，在臺灣，巨量資料的應用雖然才剛剛萌芽，但是要面臨的幾項重大挑戰，除了資料分析人才不足之外，其他還有在地顧問服務不足以及對資料價值的敏感度不足的問題，而這些因素都將影響巨量資料在臺灣市場的發展，故如何快速協助製造現場領域專家以資料分析角度有效掌握製程品質為本研究發展重點。

基於上述需求，本研究將研發具整體學習機制在製品品質檢測暨分類模組關鍵技術，在產線製程特性不明確下，利用向量機、邏輯斯迴歸及決策樹等多種機器學習機制進行建模，可達成穩健的品質分類效能，協助現場專家迅速發現製程異常，提升生產品質與效率之目的，以下將於第二章節說明各分類技術(向量機、邏輯斯迴歸及決策樹)原理及特性；及第三章節說明本研究所提出袋式(Bagging)整體學習方法，第四章節將說明7種交互作用模擬資料集與二類別及三類別實驗方式，並以雷達圖方式呈現本研究較單一分類技術更穩健之實驗結果進行討論，第五章節說明本研究之結論。

2 · 分類技術文獻探討

產品在生產過程中，包括製程參數設定、材料特性、生產設備與環境的感測數據，都是與產品品質有關的特徵，令一產品 j 的特徵向量 $\mathbf{x}_j = [x_1, x_2, \dots, x_d]^T$ ，當中 d 表示所有特徵向量之維度。該產品的品質 y_j 可透過歷史資料建立的分類模型進行品質分類，常用的分類模型包括決策樹(Decision Tree) [3]、支援向量機(Support Vector Machine ; SVM) [4]與邏輯斯迴歸(Logistic Regression) [5]，簡單說明如下。

2.1 決策樹缺陷分類與信心度評估設計

決策樹是一利用樹狀結構，呈現特徵參數對分類結果 y_j 映射關係的分類方法。Quinlan 提出 ID3 [6] 演算法，後來不斷改良為 C4.5 版本與 C5.0 商用版本 [7]，其他常見的演算法包括：CART [1]、CHAID [8] 等。決策樹從根節點由上而下，透過依序對每一特徵值尋找最佳的切點位置，把資料由父節點切分到子節點上，在每一個節點上都只利用該節點內的資料繼續往下尋找。

所謂最佳的切點位置，每個演算法採用不同的準則，一般常以基於資訊亂度的 Gini Diversity Index (GDI) 計算。GDI 表示資料的不純度，所以當一個節點內都是同一類別(例如都是品質正常，或都是有品質缺陷的產品)，亂度最小，在建立決策樹的過程，對每一特徵值尋找 GDI 最小的切點位置，計算方式如下：

$$1 - \sum_i p^2(i) \quad (1)$$

其中 $p(i)$ 是表示在一節點中，第 i 個類別資料出現的機率，通常以資料比例代表之。

類似的亂度計算方式為 Deviance (或 Entropy) 計算方式如下：

$$-\sum_i p(i) \log p(i) \quad (2)$$

其中 $p(i)$ 同樣是一節點中，第 i 個類別資料出現的機率。

另一種用則是 Twoing rule，計算左右子節點的比例差值，數值越大表示分類效果越好：

$$P(L)P(R) \left(\sum_i |L(i) - R(i)| \right)^2 \quad (3)$$

其中 $P(L)$ 與 $P(R)$ 是左右兩個子節點的機率，以資料量比率代表， $L(i)$ 與 $R(i)$ 則是第 i 個類別資料分別在左節點與右節點的數量，比較不同的是 Twoing splitting rule 是越大越好，表示同類的資料都在同一節點。

而 GDI 與 Deviance 差別不大，如下圖，以兩類資料模擬計算，兩個準則只是數值大小與變化趨勢不同，但同樣在資料最亂時(二類分的 $p=0.5$) 達到最大值，表示此時亂度最大，訓練時則是選用能達到亂度最小的節點分裂方式。

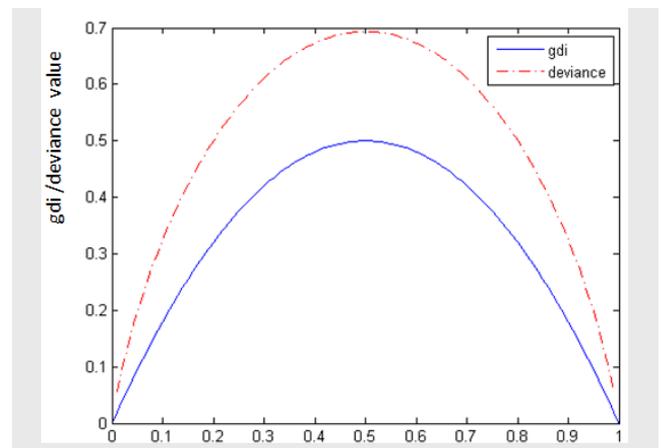


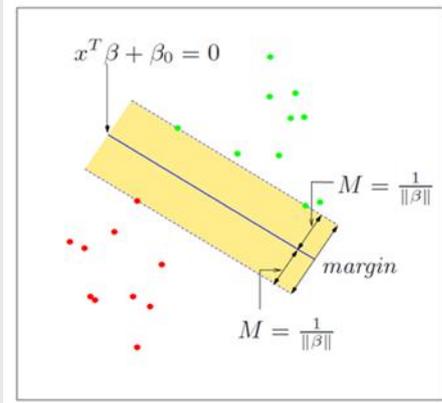
圖 1 決策樹 GDI 與 Deviance 之比較

在信心度評估方面，以每個葉節點中，判定類別所佔的資料比，作為類別判斷之信心度。

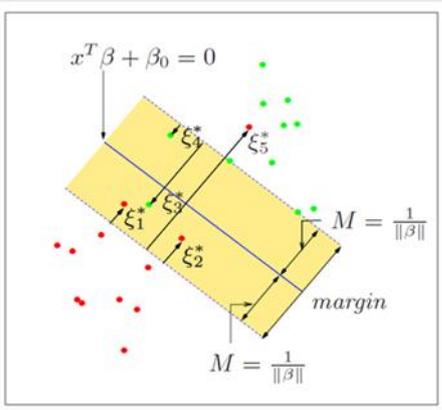
2.2 SVM 缺陷分類與信心度評估設計

傳統上，SVM 是一個基於二分類問題所設計的分類器 [9]，於 R^p 空間上建構一個高維度的超平面 $f(x) = \beta_0 + x^T \beta = 0$ 將空間一分为二。在最理想的情況下， $f(x) = 0$ 能夠將分別屬於兩類的資料完全分開(如圖 2 上)，使得超平面的兩側各只有一類資料。而能夠如此切割空間的超平面通常不會只有一個，在 SVM 方法中，我們稱超平面到距其最近訓練資料點 (Support vectors，圖中平面兩側虛線上的點) 之間的距離稱為間隔 (Margin, M)，SVM 的目標即為找出間隔最大的超平面來分別兩類資料。實務上，兩類資料點往往會交疊在一起而無法從中找到超平面去切割開的(如圖 2 下)，因此，SVM 能夠容許部分資料落於超平面「錯誤的」一側，僅要求這些資料點與超平面的距離 $\zeta_i^k = M \zeta_i$ 能夠盡量地小(若資料點落於超

平面正確的一側，則該 $\xi^* = 0$)。綜合以上考量，我們可寫下SVM求解超平面的問題即為一個最佳化問題。



(上)



(下)

圖 2 SVM示意圖:(上) 兩類資料完全分散開;(下) 兩類資料於空間中重疊

但在面對多分類問題時，需做進一步的擴展。本研究以R tool e1071 package [10]，採用SVM功能，實現SVM缺陷分類器。其中利用”One-against-one”的擴展方法，即假設資料共有 k 個類別，需要建立 $k(k-1)/2$ 個SVM分類器。每次針對兩個類別的資料，例如第 i 、 j 類，建立二分類之SVM分類器，如下所示：

$$\begin{aligned} \min_{w^{ij}, b^{ij}, \xi^{ij}} & \frac{1}{2} (w^{ij})^T w^{ij} + C \sum_t (\xi^{ij})_t \\ \text{subject to} & \\ (w^{ij})^T \phi(x_t) + b^{ij} & \geq 1 - \xi_t^{ij}, \text{ if } x_t \text{ in the } i\text{th class,} \\ (w^{ij})^T \phi(x_t) + b^{ij} & \leq -1 + \xi_t^{ij}, \text{ if } x_t \text{ in the } j\text{th class,} \\ \xi_t^{ij} & \geq 0. \end{aligned} \quad (4)$$

在判斷一測試資料 x 的類別時，採投票(voting)機制，即視 $k(k-1)/2$ 個SVM分類器的分類判斷結果，採得票數最多的類別，作為最終

判斷的分類類別。

在信心度評估方面，以判定一測試資料 x 屬於某類別時的機率，作為類別判斷之信心度。例如判斷 x 為第 i 類之信心度如下：

$$p_i = P(y = i | x), \quad i = 1, \dots, k. \quad (5)$$

因此目標是求解出SVM分類器判斷 x 為第 i 類時的機率 $p_i, i = 1, \dots, k$ 。首先由於我們採”One-against-one”的作法，根據兩兩建立的SVM分類器模型，求解 r_{ij} ：

$$\begin{aligned} r_{ij} & \approx P(y = i | y = i \text{ or } j, x) \\ & \approx \frac{1}{1 + e^{A\hat{f} + B}} \end{aligned} \quad (6)$$

其中， \hat{f} 為 x 根據該SVM模型判斷的類別，則可利用 Negative log likelihood 求出 A 、 B ，進而求出 r_{ij} 。

接著根據條件機率，可得以下等式：

$$\begin{aligned} P(y = j | y = i \text{ or } j, x) \cdot P(y = i | x) &= \\ P(y = i | y = i \text{ or } j, x) \cdot P(y = j | x) & \end{aligned} \quad (7)$$

最後將求解 p 轉換為求解以下的最佳化問題，可解得 p ：

$$\begin{aligned} \min_p & \frac{1}{2} \sum_{i=1}^k \sum_{j \neq i}^k (r_{ji} p_i - r_{ij} p_j)^2 \\ \text{subject to} & p_i \geq 0, \forall i, \sum_{i=1}^k p_i = 1. \end{aligned} \quad (8)$$

2.3 邏輯斯迴歸式缺陷分類與信心度評估設計

傳統上，邏輯斯迴歸是一個基於二分類問題所設計的分類器，其模型假設如下：

$$\log \frac{\Pr(Y=1 | X=x)}{\Pr(Y=0 | X=x)} = \beta_0 + x^T \beta \quad (9)$$

首先，這是一個機率比值(Odds Ratio)的概念，也就是假設在給定 $X = x$ (或說觀測到 x) 的前提下，其類別是屬於 $Y = 1$ 的機率，比上其類別是屬於 $Y = 0$ 的機率，取對數(log)後是一個線性模型 $(\beta_0 + x^T \beta)$ 。

據此定義， $\Pr(Y = 1 | X = x)$ 的機率值，可以用下列方程式計算：

$$\Pr(Y = 1 | X = x) = \frac{\exp(\beta_0 + x^T \beta)}{1 + \exp(\beta_0 + x^T \beta)} \quad (10)$$

在面對多分類問題時，需做進一步的擴展。而在信心度評估方面，亦以判定一測試資料屬於某類別時的機率 \Pr ，作為類別判斷之信心度。當類別有 K 種時，選第 K 類為 odds ratio

的比較基礎，可以定義以下 $K-1$ 組的 odds ratios：

$$\begin{aligned} \log \frac{\Pr(y_i = 1 | x_i)}{\Pr(y_i = K | x_i)} &= \beta_{0,1} + x^T \beta_1 \\ \log \frac{\Pr(y_i = 2 | x_i)}{\Pr(y_i = K | x_i)} &= \beta_{0,2} + x^T \beta_2 \\ &\vdots \\ \log \frac{\Pr(y_i = K-1 | x_i)}{\Pr(y_i = K | x_i)} &= \beta_{0,K-1} + x^T \beta_{K-1} \end{aligned} \quad (11)$$

經整理可得：

$$\begin{aligned} \Pr(y_i = 1 | x_i) &= \frac{\exp(\beta_{0,1} + x^T \beta_1)}{1 + \sum_{k=1}^{K-1} \exp(\beta_{0,k-1} + x^T \beta_{k-1})} \\ \Pr(y_i = 2 | x_i) &= \frac{\exp(\beta_{0,2} + x^T \beta_2)}{1 + \sum_{k=1}^{K-1} \exp(\beta_{0,k-1} + x^T \beta_{k-1})} \\ &\vdots \\ \Pr(y_i = K-1 | x_i) &= \frac{\exp(\beta_{0,K-1} + x^T \beta_{K-1})}{1 + \sum_{k=1}^{K-1} \exp(\beta_{0,k-1} + x^T \beta_{k-1})} \end{aligned} \quad (12)$$

上式中，各 β 的估計，可根據 MLE (Maximum Likelihood Estimation)，偏微分推倒求得收斂式之後可解出。

3 · Bagging式整體學習機制

由於訓練資料的蒐集，不容易蒐集到完整的母體，或是完整代表母體的資料，若是訓練分類器時過度擬合(Overfitting)，反而在真實應用的時候，會得到較差的效果。部分的分類器為了避免過度擬合，會加入懲罰項來減緩模型過度擬合於訓練資料；但懲罰項的設計，卻不容易適用於各種輸入資料的分佈。Breiman [11] 於1996年所提出Bootstrap Aggregation，對訓練資料的反覆重抽樣，分別訓練多個版本的分類器，再由分類器結果平均或投票來決定Bagging的分類結果，如下圖：

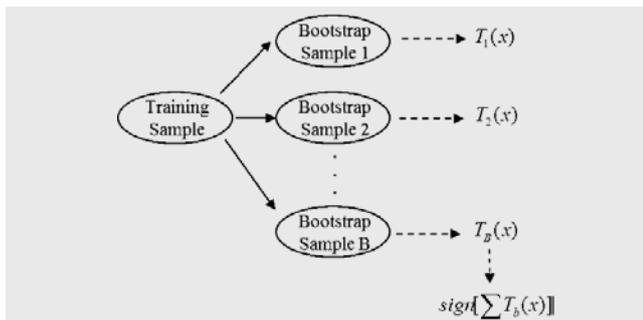


圖 3 傳統Bagging架構

Bagging藉由分類器的均方差 (Mean Square Error) 說明複合多個分類器的均方差小於或等於單一分類器的均方差。令 A 為 M 個樣本的訓練資料集：

$$\Lambda = \{(y_m, x_m), m = 1, 2, \dots, M\} \quad (13)$$

其 y_m 為分類目標， x_m 為多維度之輸入參數。

單一分類器 ψ 在輸入 x 之分類結果為 $\psi(x, \Lambda)$ ，則可定義 B 個子訓練資料集合 $\{\Lambda_b\}$ ，由 A 以 uniform 分佈投返式抽樣獲得，以及對應之分類器集合為 $\{\psi(x, \Lambda_b)\}$ 。Bagging 分類器由 $\{\psi(x, \Lambda_b)\}$ 各自分類結果做整體學習：

$$\psi_A(x) = E_\Lambda \psi(x, \Lambda) \quad (14)$$

其中 E_Λ 依據分類目標而有不同：若分類目標 y_m 為數值型，則為 $\{\psi(x, \Lambda)\}$ 的期望值，若是類別型，則為投票結果。

輸入的每個樣本為多維度，但並非能一次在所有維度觀察出該樣本的分佈特性，故解決方式常見有三種，其一是依賴分類器的能力，交由分類器去判斷各自維度對於該資料在訓練資料集中的解釋度；但此舉亦加重倚賴分類器的分類效能，若碰到分類器不適用時，則會明顯降低分類結果。其二是先做前處理，如 PCA [12]，再套入分類器；但原始資料分佈被轉換到 Eigenspace，其特性已經迥異，並不能確定適用於所有分類器。其三是混入特徵挑選，每次取樣的時候都挑選一定比例以上的維度，套入分類器之中；雖然特徵挑選較為耗時，但 Bagging 非常適合應用在雲端的平行化環境下執行，透過平行化方式則可有效降低實用性的影響。本研究中，Bagging 設計為特徵挑選後，再行套入該次分類器的訓練。

由於每個分類器的設計皆有其初始的假設，並無法適用於各種不同的原始資料分佈中。本研究設計的 Bagging 方式 (itriBagging)，引入多種不同的分類器，因應不同的原始資料分佈，得到各自的分類結果，藉此提供對於多種資料分佈的適應性。為此，也必需將不適用於當前原始資料分佈的分類器過濾掉。

下圖為本研究 itribagging 之模型訓練架構。模型定義如下：

$$\phi_B(x) = E_\Lambda(x, \Lambda, \zeta, \beta, \gamma) = \sum_{b=1}^B \gamma_b \phi(\zeta_b(x), \beta_b) \quad (15)$$

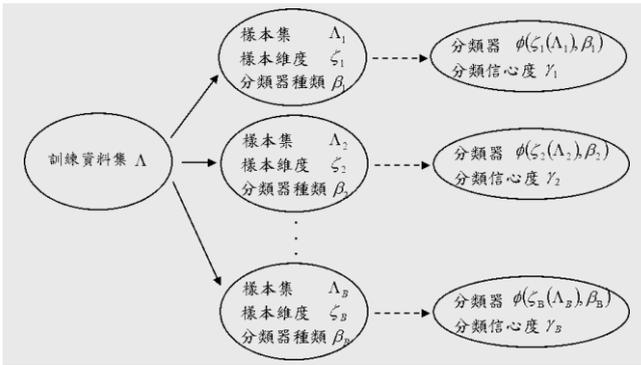


圖 4 itriBagging 統合機制模型訓練架構

包含五個資訊：建立的模型數量 B 、樣本點 x_m 在該模型中被挑選的樣本維度 $\zeta = \{\zeta_1, \zeta_2, \dots, \zeta_B\}$ 、該分類器種類 $\beta = \{\beta_1, \beta_2, \dots, \beta_B\}$ 、以及該分類器在訓練時所計算的分類信心度 $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_B\}$ 。建立每一個分類器時，依照預先設定的比例，抽取部份作為訓練，部份保留為該單一分類器之外部測試(Outside Test)，用以計算分類信心度。分類器的選擇，則使用Uniform Distribution挑選。

下圖為Bagging分類處理的架構圖。假設現有測試資料集 T

$$T = \{y_n, x_n\}, n=1, 2, \dots, N \quad (16)$$

以及bagging 模型 $\phi_B(x)$ ，則分類結果決定方式為依據各自分類器種類 β ，各模型挑選之樣本維度 ζ ，所得之分類結果，於計算平均或投票時乘上分類信心度 γ ，即為模型分類結果。

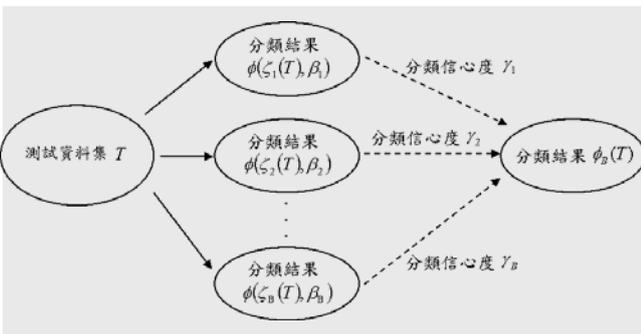


圖 5 Bagging 統合機制模型分類處理架構

以下用一個例子說明模型訓練、測試流程：假設訓練資料集中共有10,000筆資料，每筆資料的維度為100，意即 Λ 為包含10,000個樣本的訓練資料集。

首先在訓練模型時，以投返式抽樣抽出3個子訓練樣本集，每個子訓練樣本集中皆由10000筆中以投返式抽樣，取出指定比例的資料，如90%，

即 $B=3, \Lambda_1 \sim \Lambda_3$ 皆為9,000個樣本的樣本集。對第一個子樣本集 Λ_1 ，從100個資料維度中挑選80個資料維度來進行訓練，即 $\zeta_1=80$ 。子樣本集 Λ_1 中每筆資料 x 經維度挑選後，剩下80個維度，表示為 $\zeta_1(x)$ 。對此樣本集，挑選Decision Tree作為分類器，即 β_1 =Decision Tree。最後用該9,000筆挑選過維度的資料，以非投返式挑選一定比例，例如85%，共7,650筆資料，建立Decision Tree分類模型，表示為 $\phi(\zeta_1(\Lambda_1), \beta_1)$ ，並以其餘15%，共1,350筆資料，進行outside test以計算該模型之分類信心度 $\gamma_1=90\%$ 。

對第二個樣本集 Λ_2 ，做法相同，令 $\zeta_2=80$ 。需注意雖然同樣挑選80個維度，但由於是隨機挑選，故第二個樣本集所挑出的維度，與第一個樣本集所挑出的維度不盡相同。對此樣本集，令 β_2 =SVM，建立SVM分類模型，表示為 $\phi(\zeta_2(\Lambda_2), \beta_2)$ ，並計算出該模型之分類信心度 $\gamma_2=70\%$ 。

同理，對第三個樣本集 Λ_3 ，維度挑選僅受限於一定比例以上，以隨機決定數字。令 $\zeta_3=85$ 、 β_3 =Logistic Regression，建立Logistic Regression分類模型，表示為 $\phi(\zeta_3(\Lambda_3), \beta_3)$ ，並計算出該模型之分類信心度 $\gamma_3=60\%$ 。

在測試階段，對任一筆測試資料 t ，決定其分類結果之流程如下：

將 t 代入第一個分類模型，所得分類結果為 A 類，表示為 $\phi(\zeta_1(t), \beta_1)=A$ 。將 t 代入第二個分類模型，所得分類結果為 A 類，表示為 $\phi(\zeta_2(t), \beta_2)=A$ 。將 t 代入第三個分類模型，所得分類結果為 B 類，表示為 $\phi(\zeta_3(t), \beta_3)=B$ 。再參考模型的分類信心度，計算 t 為 A 類的投票分數 $= (\gamma_1 + \gamma_2)/2 = (90\% + 70\%)/2 = 80\%$ ， t 為 B 類的投票分數為 60% ，故最後決定此筆資料 t 之分類結果為 A 類。

4 · 技術驗證

4.1 資料與實驗設計

本研究以機率模型進行資料模擬，其資料產生模型主要的公式如下：

$$Y_i \sim \text{Bernoulli}(L(x_i^T \beta)), i=1, \dots, n \quad (17)$$

其中 $L(z) = (1 + e^{-z})^{-1}$ ，每個模型中，共有10個變數，設定3個為重要變數（係數不為0），第一

個係數為截距項，截距項的角色在於控制 $Y_i=1$ 的比率，截距項的係數愈大 $Y_i=1$ 的比率愈高，而我們設定 $Y_i=1$ 為產品有缺陷， $Y_i=0$ 為產品正常。靠調整資料產生函數的線性部份來產生 15 種不同交互特性的資料，以下將以 7 種代表性資料模型進行實驗結果討論，包括 Interaction (Model_6)、Jump function (Model_8)、Sine function (Model_10)、Correlation (Model_12)、Random (Model_13)、Autoregressive (Model_14)、Multi-staged (Model_15) 模型。由於 SVM 與 LR 分類器常用於二類別分析，故本研究除了以二類別來驗證效能外，另以三類別資料來進行多類別之分類實驗，除了實作各模型多類別模擬資料的產生，也同時驗證相關分類器對於各種模型下多類別資料的分類效果。

實驗的步驟如下圖 6 所示，首先產生 1,000 組模擬資料，每一組模擬資料皆事先分好訓練組與測試組，四種分類器皆以相同的訓練組來訓練分類模型，再使用分類模型對測試組進行分類，根據模型分類結果與測試資料實際值，便能得知分類模型的分類準確度。

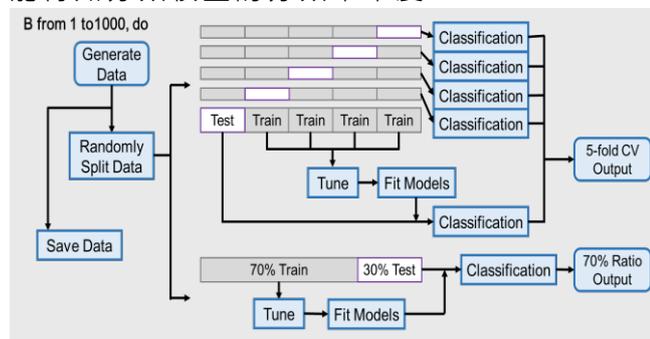


圖 6 產生模擬資料與建立模型流程

測試組與訓練組的安排分為兩種（如上圖），其一採取 K -fold 交叉驗證的方式 [14] 將資料切為五等份 ($K=5$)，依序以其中四份訓練模型，進行所剩一份資料的分類實驗。最後整合五組測試組的分類結果與實際值比對，檢視分類準確度。其二為依比例來抽取，固定抽取 70% 資料來訓練模型，其餘 30% 資料作為測試組。

實驗用的分類器為 LR、SVM、Decision Tree (實驗結果簡稱 Tree) 及本研究所開發的 itriBagging，為求適合訓練組資料的最佳模型，我們在每次訓練分類模型時，皆給以一定範圍的模型參數進行模型調校，如表 2。SVM 的模型參數為控制誤判情形的可容許範圍 C

(cost)；TREE 的模型參數有兩個，minsplit 為一樹節點在切割時至少有的觀測值數目，minbucket 為一終端節點至少有的觀測值數目，而 itriBagging 可調整重抽的次數、參數及資料抽出的比例。

實驗使用正確率 (Accuracy, ACC)、真陽性率 (True Positive Rate, TPR) 與真陰性率 (True Negative Rate, TNR) 來評估分類模型分類準確度。ACC 為模型分類結果與實際值一致的比例；TPR 為實際上有缺陷的資料被分類模型準確分類的比例；TNR 為實際上正常的資料被分類模型準確分類的比例。

表 2 各分類器實驗參數設定範圍

分類模型	參數	範圍
SVM	C (cost)	$\{2^k \mid -4 \leq k \leq 4\}$
Decision Tree	Minsplit	{10, 20, 30}
	Minbucket	{20%, 30%, 40%}
itriBagging	Bagging 重抽次數	{50, 100}
	Feature 抽出比例	{70%, 90%, 100%}
	Data 抽出比例	{70%, 90%, 100%, 120%}

4.2 實驗結果

4.2.1 二類別分類實驗結果

(1) 不同交叉驗證法 (Ratio 70% 及 5-fold) 的結果並無顯著差異，此結果是合理的，因為樣本數夠大 ($n=400$) 且模擬的次數夠多 (1,000 次)，且因為此二種交叉驗證法都是不偏的，所以理當算出極為類似的結果，這裡我們選用依比例抽取的資料來繪製結果。

(2) 如圖 7-9，考慮交互作用時 (Model_6)，TREE 及 itriBagging 表現較佳，但對於其它模型，TREE 的表現比則略遜一籌。

(3) 各分類器對於 Model_8、10、12 分類結果較佳，而對於 Model_13 ~ Model_15 分類結果略差一些，可顯現出不同特性的模型分類難度不同。

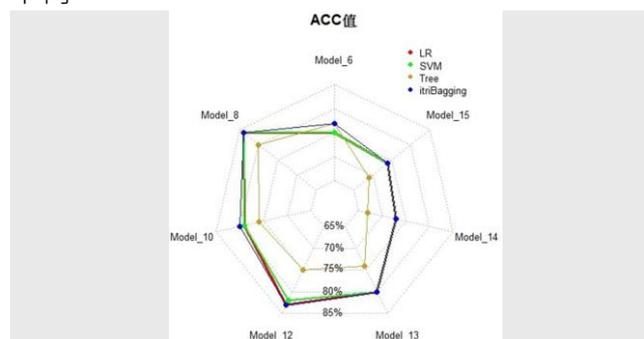


圖 7 二類別分類結果 (ACC)

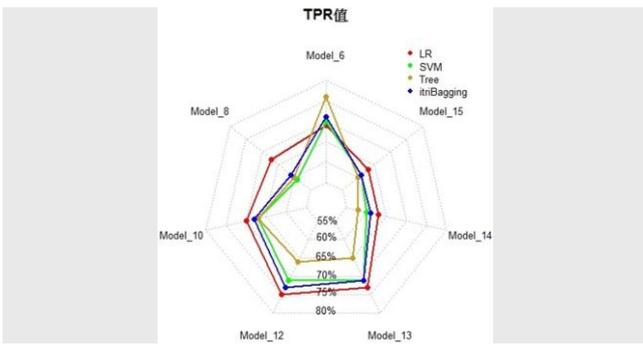


圖 8 二類別分類結果 (TPR)

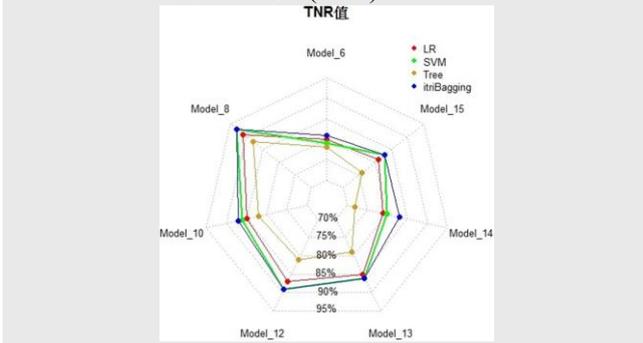


圖 9 二類別分類結果 (TNR)

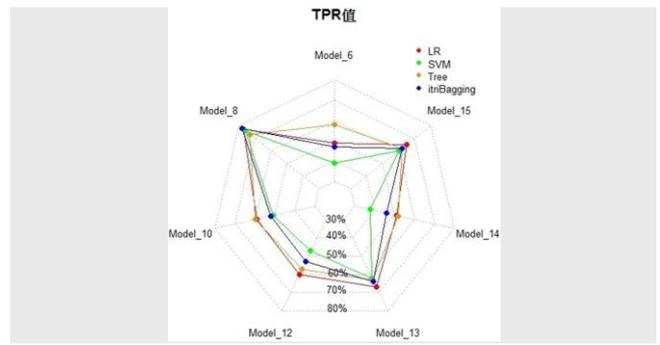


圖 11 三類別分類結果 (TPR)

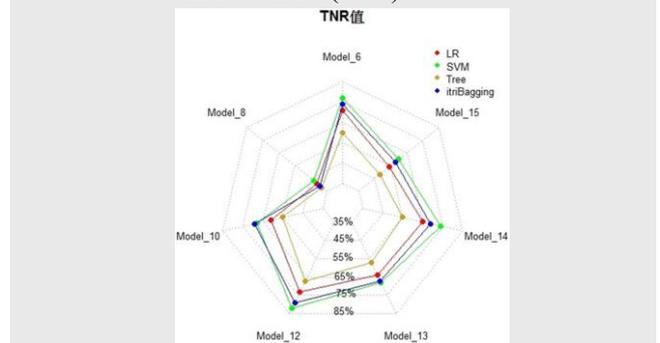


圖 12 三類別分類結果 (TNR)

4.2.2 三類別分類實驗結果

(1) 在三類分類結果分析中，TPR與TNR在並無法明確的定義，在此，我們把三類的結果分成兩類：「reference」（ $Y=0$ ）及「非reference」（ $Y=1、2$ ），然後再套用二分類的TPR與TNR計算公式得到最後結果。

(2) 如圖10-12，各分類器對於Model_6、8、10、12分類結果較佳，而對於Model_13及Model_15分類結果略差一些，可顯現出不同特性的模型分類難度不同，以ACC及TNR來看，itriBagging的表現優於其它分類器，以TPR來看，各種分類器對於不同模型各有所長。

(3) 三分類之ACC遠低於二分類之ACC，這個結論合理，因為三分類模型的參數是二分類的兩倍，所以用相同的樣本數來訓練模型會比較不準確。

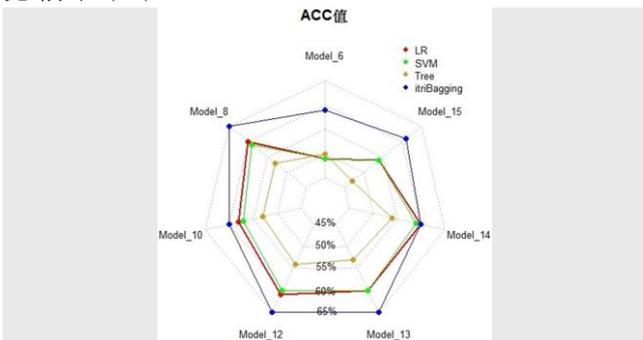


圖 10 三類別分類結果 (ACC)

4.2.3 驗證結果討論

(1) 以二類別而言，itriBagging與LR的ACC相仿，但是LR有較高的TPR而itriBagging有較高的TNR。本研究僅根據第一組資料來挑選itriBagging的參數組，未來可進行最佳參數訓練，預期可得到更好的結果。

(2) 在二類別資料之下，Tree分類器在模型有交互作用項的情形下有較好的表現，在自變數眾多時，一般LR及SVM無法考慮所有交互作用項，但是Tree本身就可以偵測到，不用強制加入交互作用項。然而這個優勢卻在三類別資料之下看不出來，由於文獻上對Tree應用於三類別資料的結果描述較少，可列為未來研究的項目。

(3) 在三類別資料之下，itriBagging表現最佳，但其所需的計算量較高於其他分類器。

5. 結論

本研究所發展之整體學習機制在製品品質檢測暨分類技術，研發 Logistic Regression、Decision Tree、與SVM-based等多種缺陷分類器模組，搭配創新之單一分類器統合機制形成多重分類技術，用以補強單一分類技術無法適用於工廠資料之多元且未知特性的缺點，透過導

入不同的分類技術並發展創新統合機制後，可提供更穩健的分類效果，處理更多參數種類，並可持續擴充更多分類技術、提升分類正確率。目前亦繼續藉由不同的分類技術發展多樣成因統合機制，可檢出更多成因參數種類、讓缺陷校正決策規則更完善、提升產品良率。相關技術成果已運用於鋼鐵廠協助領域專家與現場參考相關之操作資訊，預估可降低該廠於某鑄造缺陷所造成之處理成本損失(約上仟萬元)，後續除持續提升鋼鐵業生產良率外，亦與紡織業者進行染程1次對色率不足成因分析，預估提升3.0%對色率可月節省25萬處理成本，及與封測業者合作低溫測試一次成功率差異成因分析，以提升國內製造業產品之品質，使業者有機會攻進高值化產品市場。

參考文獻

- [1] 劉郁珉, 陳蘊彥, 程瑞曦; “邁向工業4.0及生產力4.0--臺灣製造業的挑戰與機會,” *電機月刊*, vol. 299, pp. 64-77, Nov. 2015.
- [2] 鄭芳田, “全自動虛擬量測,” *科學發展*, vol. 490, pp. 48-54, Oct. 2013.
- [3] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, CRC press, 1984.
- [4] C. Cortes and V. Vladimir, “Support-vector networks,” *Machine learning*, vol. 20(3), pp. 273-297, Sep. 1995.
- [5] D. W. Hosmer Jr and S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed., Wiley press, 2013.
- [6] J. R. Quinlan, “Induction of decision trees,” *Machine learning*, vol. 1(1), pp. 81-106, Mar. 1986.
- [7] J. R. Quinlan, *C4. 5: Programs for Machine Learning*, Elsevier press, 2014.
- [8] G.V. Kass, “An exploratory technique for investigating large quantities of categorical data,” *Applied statistics*, vol. 29(2), pp. 119-127, 1980.
- [9] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd

ed., New York, Springer, 2009.

- [10] D. Meyer et al. (2015) R Tool package e1071 [online]. Available: <https://cran.r-project.org/web/packages/e1071/index.html>
- [11] Leo Breiman, “Bagging Predictors,” *Machine learning*, vol. 24(2), pp. 123-140, Aug. 1996.
- [12] K. Pearson, “On Lines and Planes of Closest Fit to Systems of Points in Space,” *Philosophical Magazine*, vol. 2(11), pp. 559-572, 1901.
- [13] Y. Freund and R. E. Schapire, “Experiments with a New Boosting Algorithm,” in *Proc. 13th Int. Conf. Machine Learning*, 1996, pp. 148-156.
- [14] S. Geisser, *Predictive Inference: An Introduction*, New York, NY: Chapman and Hall, CRC press, 1993.

作者簡介

鄭光宏



現任工研院資通所/智能製造服務系統組資料應用部工程師，專長為機器學習、資料探勘等

夏啟峻



現任工研院資通所/智能製造服務系統組資料應用部工程師，專長為資料探勘、多變量統計分析、圖樣識別、實驗設計、訊號處理等

葉峻賓



現任工研院資通所/智能製造服務系統組資料應用部工程師，專長為機器學習、圖樣識別、訊號處理、資料探勘等

江宜霖



現任工研院資通所/智能製造服務系統組資料應用部工程師，專長為嵌入式系統、視訊編碼、機器學習等

林順傑



現任工研院資通所/智能製造服務系統組資料應用部研發副理，專長為機器學習、訊號處理、數據分析、資料探勘等

蔡煥文



現任工研院資通所/智能製造服務系統組資料應用部經理，專長為資料探勘、機器學習、統計分析等