

資料探勘技術在晶圓針測誤宰分析之應用

Applying Data Mining Techniques to the
Overkill Analysis of Wafer Testing

學 生：丁世杰

Student : Shih-Chieh Ting

指導教授：劉敦仁 博士

Advisor : Dr. Duen-Ren Liu



Submitted to Degree Program of Applied Science and Technology

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master in

Degree Program of Applied Science and Technology

July 2012

Hsinchu, Taiwan, Republic of China

中華民國一零一年七月

國立交通大學理學院應用科技學程碩士班

學生：丁世杰

指導教授：劉敦仁 博士

國立交通大學
理學院應用科技學程

摘 要

半導體晶圓針測(Wafer Probing)為半導體晶圓製造完成後，對產品良率進行驗證的重要程序，其目的在於找出晶圓上晶粒的瑕疵；主要以探針的方式與每顆晶粒上的鐳墊接觸，輸入電流或信號，並接收所反饋出來的數值，以判斷是否為良品。然而，由於測試時可能因為測試機台、針測機或探針卡的穩定性不足、操作人員的參數設定動作不確實，而造成將良品的晶粒誤判為壞品的誤宰(Overkill)現象出現；此現象會影響產品的良率，事後還必須進行重測程序，而使生產效率降低。對半導體晶圓測試產業來說，誤宰現象除了造成測試成本的浪費，也嚴重影響客戶的信任度。

本研究透過與產業領域專家的溝通討論後，建立多項與誤宰現象有相關的偵測演算法，並經由資料探勘分類技術 (Classification)，針對這些偵測演算法所計算結果，結合實際的晶圓針測資料為實驗數據進行分析，對這些偵測演算法的準確度進行確認，以用來協助業界在實際生產時參考，確認是否為誤宰的決策輔助。

由於造成誤宰的根本原因，包含了設備、針測卡損耗、人員設定問題，甚至是這些狀況的組合。本實驗也從實際的針測資料中，透過運用資料探勘技術的關聯規則 (Association Rule)，找出與晶圓誤宰有關的可能原因，做為製程改善的參考。

關鍵字：晶圓針測、誤宰、資料探勘

Applying Data Mining Techniques to the Overkill Analysis of Wafer Testing

Student : Shih-Chieh Ting

Advisor : Dr. Duen-Ren Liu

Degree Program of Applied Science and Technology

College of Science

National Chiao Tung University

Abstract

Wafer probing is a critical process employed to measure the yield of wafer fabrication. The major object of wafer probing is to find the defect dice on the wafer. It inputs electrical current and signals through the probe needles which contact the pad of each dice and receives the outputs to determine whether or not the dice are good.

However, the probing result could be affected by the stability of tester, prober, probe card or the setting actions of operators. Overkill situations happen if good dice are misjudged as bad dice caused by one or more of the above factors causing an abnormal yield and requiring re-probe actions which diminish production performance and the trust of customers.

In this study, we talked to the specialists of this industry in order to build some overkill-related detection methods. The aim was to implement these detection methods based on the real wafer probing data from one of Taiwan's testing facilities. We used classification technologies of data mining on those results generated by these detection methods to determine the correctness of these methods, to ascertain if they could be implemented in the data analysis system as a real-time alarm to handle the overkill issue.

The root causes of overkill situations are not easy to find. They include: stability of tester, usage of probe card, setting actions by operators or a combination of these factors. We used association rule technology of data mining to find the possible causes of overkill situation, to serve as reference for improving actions.

Keyword: Wafer Probe, Overkill, Data Mining

誌 謝

回顧轉眼即逝的兩年在職專班生活，既忙碌又充實，收獲豐富。

首先感謝指導老師劉敦仁教授的費心指導，在題目以及實驗方法的選取、實驗結果的分析以及論文撰寫，不遺餘力的多次協助，使本論文可以順利完成；也感謝同學俞樺、榮助及凱心，在論文寫作及編排方面給予多次的協助。

當初本著回到學校重新學習以提振工作動力的心態，確實在理學院專班得以實現，讓我重新燃起學習的熱情；要感謝學長永安，很高興二年前接受您的建議，讓我能加入交大理學院在職專班。在二年的學生生涯中，要感謝專班主任陳永富老師，孜孜不倦地導正學生的學習心態，也以幽默風趣的教學方式，讓學生在面對艱深的理論學問時，能以輕鬆的心情學習吸收；也要感謝專班同學其涎，在專班的各項學務中，替大家處理了很多繁複的工作，使專班同學都能專心於學習及研究。

感謝公司同仁在工作實務方面的知識提供，使本論文的建立可以有相關領域技術的充份支援，在我遇到無法前進的障礙時，都能順利的解決。也感謝家人的支持及鼓勵，配合我的學業調整各項活動，陪伴著我，給予打氣及愛，謝謝你們。

很高興二年前報考專班的決定，讓我的生活有了很大的改變，也朝著更好更完善的人生前進；謹以本篇論文，獻給所有給予我幫助的人。

謹誌

世杰 July 2012 新竹

目 錄

摘 要.....	I
ABSTRACT.....	II
誌 謝.....	III
目 錄.....	IV
圖目錄.....	VI
表目錄.....	VIII
第 1 章 緒論.....	1
1.1 研究動機與目的.....	1
1.2 研究範圍.....	3
1.3 論文架構.....	3
1.4 研究步驟與方法.....	4
第 2 章 文獻探討.....	6
2.1 晶圓測試簡介.....	6
2.2 晶圓測試的誤宰(OVERKILL).....	9
2.3 資料探勘.....	15
2.3.1 資料分類.....	17
2.3.2 關聯法則.....	20
第 3 章 研究架構設計.....	22
3.1 現況分析.....	23
3.2 問題定義.....	24
3.3 誤宰偵測演算法的建立.....	25
3.4 資料萃取.....	27
3.5 資料探勘.....	28
第 4 章 實驗分析.....	29

4.1	實驗一：針測誤宰決策規則建立	29
4.1.1	資料準備	29
4.1.2	資料處理	30
4.1.3	資料探勘	31
4.2	實驗二：針測誤宰原因分析	35
4.2.1	問題定義	35
4.2.2	資料準備	35
4.2.3	資料探勘	38
第 5 章 結論與未來研究方向		41
參考文獻		43
附 錄		45



圖目錄

圖 1：研究流程	5
圖 2：晶圓測試之探針、電性接點、晶粒、晶圓之關係示意圖	6
圖 3：晶圓測試項目範例	7
圖 4：晶圓測試 MAP 圖形	8
圖 5：測試結果摘要	8
圖 6：晶圓定位示意圖	9
圖 7：顯微鏡下的晶圓定位	10
圖 8：良好的探針痕跡	10
圖 9：異常的探針痕跡	10
圖 10：針壓過大導致底部露出	11
圖 11：多並測數探針未確實接觸的痕跡	11
圖 12：晶圓誤宰原因魚骨圖	12
圖 13：異常的晶圓針測圖形	13
圖 14：單一測試 SITE 的良率過低	13
圖 15：與測試移動路徑類似的連續 DEFECT	14
圖 16：測試 MAP 平移(人員設定問題)	15
圖 17：多個超平面可切開類別，但只有一個能達到最大分割	18
圖 18：決策樹基本架構	19
圖 19：系統資料探勘架構	22
圖 20：測試資料傳輸架構圖	23
圖 21：特定排列方式的 BIN MAP	25
圖 22：晶圓測試結果摘要	26
圖 23：資料處理範例	27
圖 24：額外增加的 OVERKILL 欄位	30
圖 25：誤宰偵測演算法的計算結果	31
圖 26：A 產品決策樹規則	33
圖 27：B 產品決策樹規則	34

圖 28：資料萃取範例.....	36
圖 29：準備進行關聯規則分析的資料欄位.....	37
圖 30：APRIORI 參數設定.....	38
圖 31：關聯規則分析結果.....	39
圖 32：問題機台每週誤宰筆數統計圖.....	40
附錄圖 1：A 產品 NAïVE BAYES 演算法執行結果.....	45
附錄圖 2：B 產品 NAïVE BAYES 演算法執行結果.....	45
附錄圖 3：A 產品 DECISION TREE(J48)演算法執行結果.....	46
附錄圖 4：B 產品 DECISION TREE(J48)演算法執行結果.....	46
附錄圖 5：A 產品 MULTILAYERPERCEPTRON 演算法執行結果.....	47
附錄圖 6：B 產品 MULTILAYERPERCEPTRON 演算法執行結果.....	47
附錄圖 7：A 產品 SVM(SMO)演算法執行結果.....	48
附錄圖 8：B 產品 SVM(SMO)演算法執行結果.....	48



表目錄

表 1：資料欄位說明.....	27
表 2：資料探勘技術與 WEKA 演算法對照表.....	28
表 3：各演算法的參數設定.....	32
表 4：使用不同分類方法檢驗偵測演算法的準確度.....	33
表 5：誤宰原因分析用資料欄位.....	36



第 1 章 緒論

台灣半導體製造產業在經過多年的成長後，為追求成本的降低及更高的毛利，紛紛進行更高度的專業分工，將資源集中到前段製程設備的投資上，而將後段晶圓測試及封裝部份製程，交由專業測試廠及封裝廠進行。由於多年來的分工經驗，晶圓測試產業已成為一成熟的產業，幾乎只要投資購置測試機台，便可以進行接單生產；因此，在面對同業的競爭下，必須提昇測試技術及生產管理能力，才能建立起獨特的競爭優勢。

對晶圓測試業者來說，測試出來的資料就是生產的產品，產品品質指的就是其資料是否正確？能否在低時間成本的前提下，正確的測試出晶圓上各晶粒的良率，就代表著業者的競爭力；而任何誤宰（將良品晶粒誤判為壞品）的狀況，不僅必須重新測試而對造成測試成本的增加，同時也直接影響到其獲利。

本論文以晶圓測試業者為研究目標，運用資料探勘技術，對晶圓測試的誤宰現象進行判斷規則的建立以及原因的分析，希望協助業者能在誤宰現象出現時立即發現，期許在生產管理方面，節省不必要的成本耗費，以提昇其競爭力。

1.1 研究動機與目的

半導體晶圓測試時，從測試機台到晶粒之間，需經過測試機台、針測機 (Prober)、探針卡及各種傳輸界面等硬體，並經過人員操作設定各項參數；一旦其中有任何硬體穩定性不足、操作程序不正確或有硬體搭配上的累積誤差時，便容易發生將良品誤測為壞品的誤宰 (Overkill) 狀況。

當誤宰現象出現而導致針測良率低落時，必須仰賴具有經驗的工程師，對針測結果加以分析，進行實驗性的驗證、甚至換到不同機台上做交互比對，才能判斷是否為誤宰，以及是否要進行重測；此一過程不但需求高規人力，也需要花費冗長時間及測試成本。

若能對晶圓針測結果，進行與誤宰現象有相關誤宰偵測演算法的檢查，之後針對檢查結果以及誤宰資料間的關係，進行分析以找出其相關性，日後在出現特定模式的針測結果時，便可據以參考，以協助判斷是否為誤宰現象。

以往業界通常使用“統計製程控制”(SPC: Statistic Process Control)來分析針測結果，然而以統計手法所找出的故障因子，常常只是統計上的關聯意義，而非真正的原因；並且需要透過大量資料數據的計算，以及專業人員對結果的分析判斷，不易將之納入電腦系統，由系統進行自動的判斷。

和統計手法相比，資料探勘技術在近年來蓬勃發展，成為用於資料分析的重要方法。由於製程流程的漸趨複雜，工程師往往不易從大量資料中找出誤宰的實際原因，以致於更多良率的損失及生產成本的浪費；資料探勘技術能夠從大量資料中，找出未知或隱含的資訊與知識，以提供決策的參考，可以協助工程師快速而有效率地找出誤宰原因。

因此，本研究試圖以資料探勘之技術，運用於分析誤宰現象相關的特定模式，主要方向有二：

- (1) 透過與領域專家的溝通討論，建立一套針對誤宰現象的特徵進行相關運算的偵測演算法，然後再使用分類技術 (Classification) 中的決策樹推論 (Decision Tree Induction) [10]、Naïve Bayes [3]、SVM [5] 以及倒傳遞神經網路 (Back-Propagation Neural network) [13] 學習法，針對這些偵測演算法所計算出來的結果進行分析，以驗證其準確度是否足夠。並嘗試建立出一套判斷規則，應用在有異常良率發生時，作為判斷是否為誤宰現象的決策輔助。
- (2) 透過分析與針測結果相關的資料，分析誤宰時所使用的測試機台、針測機 (Prober)、探針卡、測試程式及操作人員等等參數，找出與誤宰現象間的相關性，以提供為進行改善措施的參考依據。

1.2 研究範圍

本論文先針對資料探勘技術中的分類技術 (Classification) 以及關聯規則 (Association rules)，就其相關文獻及論述，作完整的分析及探討。再逐步探討晶圓針測誤宰現象的發生原因，以及在針測資料上所呈現出的特徵，並與領域專家討論後，找出可能的變異參數，作為實驗時的分析目標。

在實驗方面，本研究以個案公司的實際晶圓針測資料庫 (Testing Parameter Analysis System: TPAS) 為資料來源，擷取正常以及異常良率的晶圓資料，運用本研究與領域專家所討論出的誤宰偵測演算法進行計算，然後以資料探勘技術中的分類技術，對這些演算法所計算的結果進行分析，以驗證其準確度。

第二個實驗，則針對這些發生異常良率時的各項環境以及設定參數，以關聯規則學習法進行分析，以期能找出與誤宰現象相關的可能原因，提供給業界作為製程與管理上的改善依據。

1.3 論文架構

本論文共分六章，各章主要內容如下：

第 1 章：緒論

說明本論文研究動機與目的、研究範圍、論文架構以及研究步驟與方法。

第 2 章：文獻回顧與探討

主要分為兩大部份；首先針對資料探勘技術的文獻進行研討，先將各種資料探勘技術的方法及工具做歸納整理，然後針對資料分類的演算法進行探討。第二部份則針對晶圓測試的誤宰現象，與其發生原因及特徵做分析。

第 3 章：研究架構設計

主要包括與個案公司的領域專家訪談所蒐集到的問題內容，以及實務上所用來判斷誤宰現象的方法，試圖歸納出自動檢查的誤宰偵測演算法；並與個案公司



原有系統配合，思考如何建立決策輔助架構，以作為下一階段實作與實驗評估之基礎。

第 4 章：實驗分析

實驗一說明如何建置誤宰偵測演算法，以及系統實作的範例。並擷取案例實際測試資料，進行數種分類技術的分析，並與實際結果進行比對，將本論文所使用的方法及結果進行探討。

實驗二則針對發生誤宰現象的針測結果，進行關聯規則的分析，以期從中找出可能的根本原因，以提供業者進行製程的改善用。

第 5 章：結論與未來研究方向

將本論文所驗證的資料探勘方法之有效度，作一完整的記錄。未來研究方向，則期望能將此架構與模型，直接套用於生產流程中，並針對誤宰偵測演算法加以改良，以期能得到更即時、更高準確度的決策輔助資訊。

1.4 研究步驟與方法

本研究的步驟為發掘問題以確認研究動機，界定出研究目的及範圍，接下來進行相關文獻探討及研讀，並與業界工程師討論，取得專業領域的相關資料，選擇資料探勘技術；確立研究方法並實地收集晶圓測試資料，然後進行實驗數據整理與分析，再歸納分析結果。研究步驟如圖 1 所示。

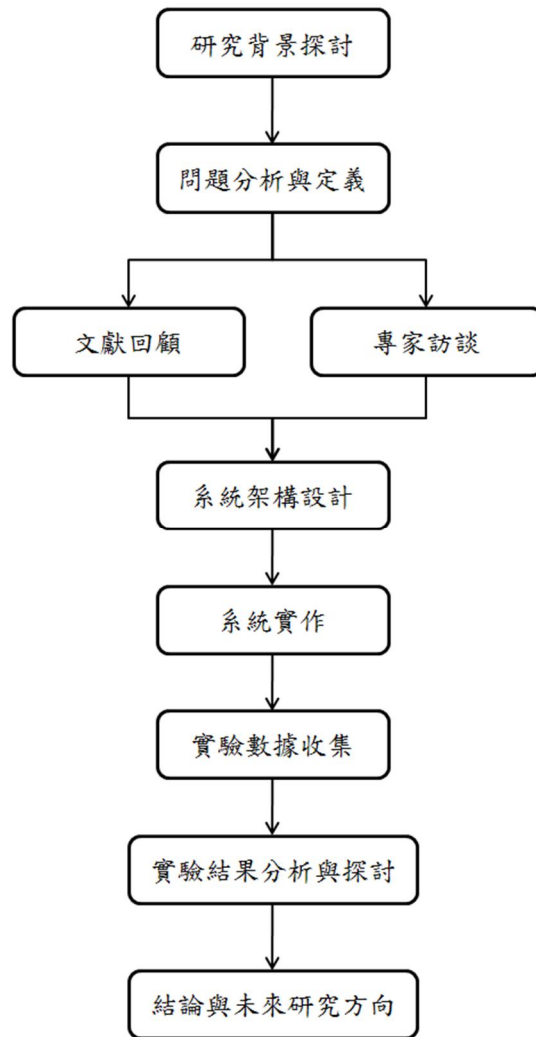


圖 1：研究流程

第 2 章 文獻探討

2.1 晶圓測試簡介

半導體是經由晶圓製造廠將空白之晶片，以照相製版的方式將先前製作好之光罩圖樣 (Pattern) 轉印到晶片上以製作光阻，再以氧化、擴散、化學蝕刻、離子植入、光阻去除與清洗...等方法，將電路逐步製作在晶片上。

完成了晶圓製作程序後，進行下一步之 IC 封裝前，為了避免封裝材料及後段設備產能的浪費，一般會進行所謂的晶圓針測 (Wafer Probe) 以將不良之晶粒 (Bad Die) 事先排除，此一流程也常被稱為晶圓分類 (Wafer sort)；意指將各晶粒分為良品與壞品。

晶圓針測的方法是利用極為精密的探針 (Probe Needle)，與晶片的電性接點 (Pad) 接觸，經由探針界面卡 (Probe Card)，將測試所需之訊號由測試機 (Tester) 送進晶粒內，並接收對應之晶粒輸出訊號，再傳回測試機台，並依測試程式來判斷晶圓功能是否符合設計規格，其關係如圖 2 所示。

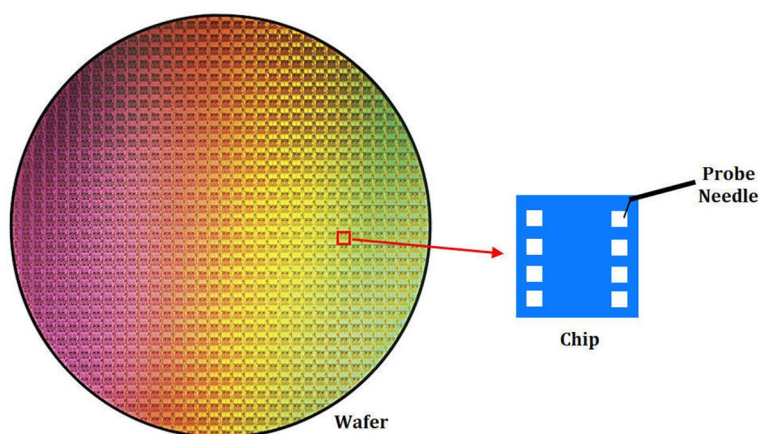


圖 2：晶圓測試之探針、電性接點、晶粒、晶圓之關係示意圖

晶圓測試 (Wafer Probe；也常稱為 Chip Probe)，是針對晶圓上的每一顆晶片 (Chip) 進行電性功能的測試，用以確認其功能是否正常。通常由客戶針對產品特

性，設計許多的測試項目，而每個測試項目則以一個測試編號(Bin Number)來代表；通過所有測試項目的晶片才能被視為是功能正常，任何一個測試項目無法通過，則該晶片就會被視為不良(Bad Die)；並在測試結果中，標示該不良晶片的失效編號(Fail Bin Number)。

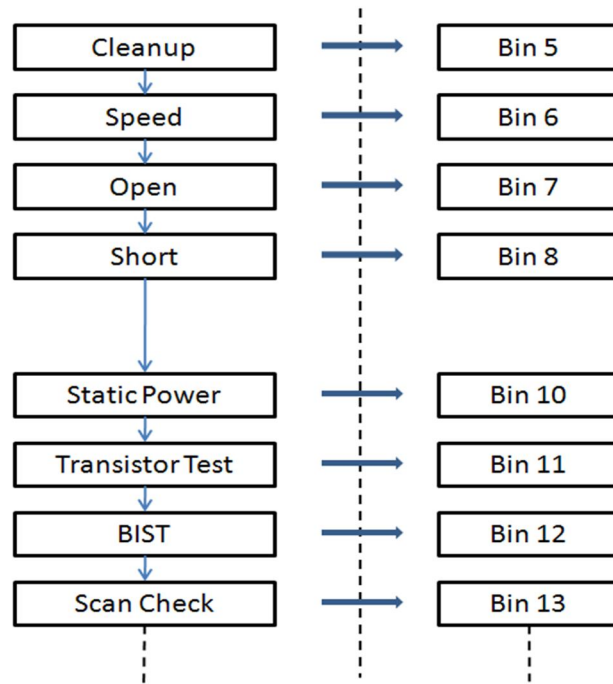


圖 3：晶圓測試項目範例

晶圓測試的結果，通常以幾何分佈的 map 圖形以及針測結果摘要(Bin Summary)呈現（如圖 4，5 所示），以利人員判斷是否有測試異常的情形以及後續資料的比對。它標示出晶圓上面的不良晶粒(Bad Die)，以及不良晶粒是因為那一個測試項目而造成不良：

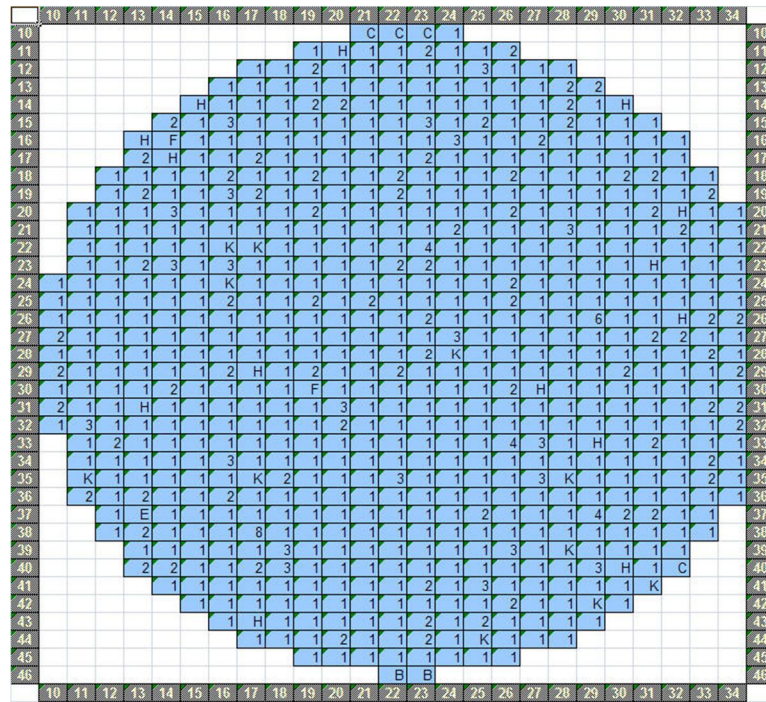


圖 4：晶圓測試 map 圖形

BIN#	CATEGORY	COUNT	RATE	S0A	S1A	S2A	S3A	S4A	S5A	S6A	S7A
1	1 Good	587	80.74%	69	75	73	78	80	73	70	69
2	2 Repair	80	11.0%	8	7	8	16	16	8	11	6
3	3 Good	21	2.89%	2	2	2	3	1	7	1	3
4	4 Repair	3	0.41%	0	0	1	0	1	0	0	1
6	6 Repair	1	0.14%	0	0	0	0	0	0	0	1
8	8 Good	1	0.14%	0	0	0	0	0	0	1	0
11	B CONTINUITY	2	0.28%	1	1	0	0	0	0	0	0
12	C LEAKAGE	4	0.55%	1	1	1	0	0	1	0	0
14	E DC	1	0.14%	0	0	1	0	0	0	0	0
15	F Repair	2	0.28%	0	0	0	1	0	1	0	0
17	H Repair	14	1.93%	1	1	5	2	1	2	1	1
20	K Unrep	11	1.51%	1	2	3	0	1	1	3	0
TOTAL		727		83	89	94	100	100	93	87	81

圖 5：測試結果摘要

此一階段所測出的良品晶粒顆數，除以整片晶圓上的晶粒顆數，便稱為晶圓測試的良率(Yield)：

$$\text{Yield} = \frac{\text{Good die count}}{\text{Total die count}}$$

從公式中可以推知最高的良率是 100%，實務上則極難達到。

通過晶圓測試之良品 (Good Die) 便可進行下一階段的封裝程序，而不合格的晶粒將會被標上記號 (Ink Dot) 並於下一個製程、晶片切割後予以丟棄。

2.2 晶圓測試的誤率(Overkill)

在晶圓測試時會使用探針卡 (Probe Card) 與晶圓上的電性接點接觸，其定位動作是由針測機 (Prober) 來控制。晶圓先透過晶圓搬運器從卡匣中將晶圓移出，並以真空吸盤固定住，然後利用光學定位系統，將晶圓移動到探針卡下方正確的位置，此時的定位包含 X-Y 平面的對準、 θ 角度的旋轉以及 Z 軸的壓力調整。

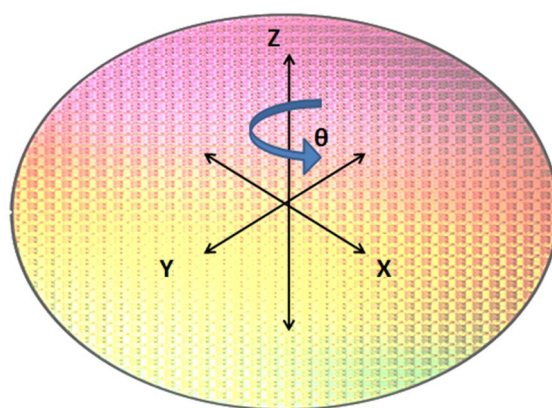


圖 6：晶圓定位示意圖

針測台上 X, Y 平面與 θ 角度的定位，會影響探針是否能正確的接觸在接點上，通常會以光學系統或電子顯微鏡來輔助定位。

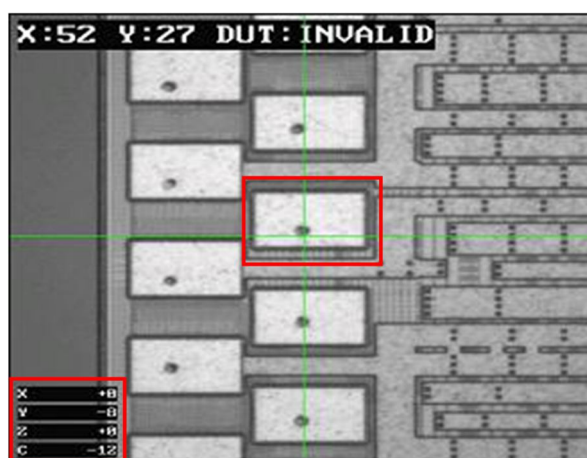


圖 7：顯微鏡下的晶圓定位

若此定位動作不準確或操作不當，探針便有可能戳到接點以外的部份，造成線路斷裂或漏電的情形，影響晶圓測試結果的正確性；圖 9 所示即為探針接觸異常所造成的痕跡。

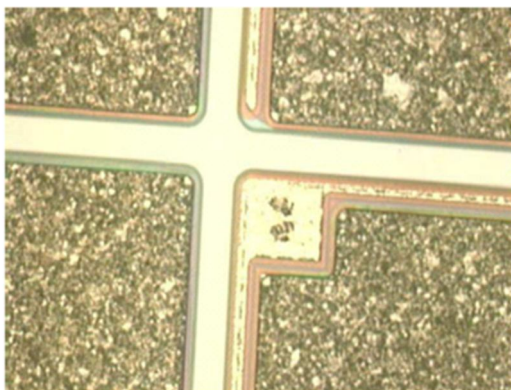


圖 8：良好的探針痕跡

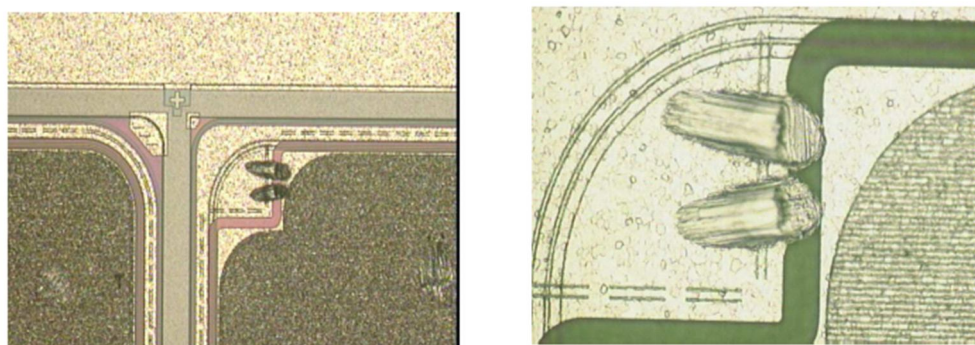


圖 9：異常的探針痕跡

針測台上 Z 軸的定位是很重要的一個步驟，它會將晶圓以垂直方向移動，使晶圓與探針卡上的探針接觸。假如移動高度太高，則造成針壓過大導致晶圓及探針有損壞的風險 (圖 10)；反之若移動高度太低，則可能造成接觸不良或未接觸；對於多個並測數(multi-site：可同時與多個晶片的電性接點接觸進行測試)的探針卡來說，由於每根探針的長度及高度可能不同，針壓過低時會可能會有一部份探針接觸不確實(圖 11)。

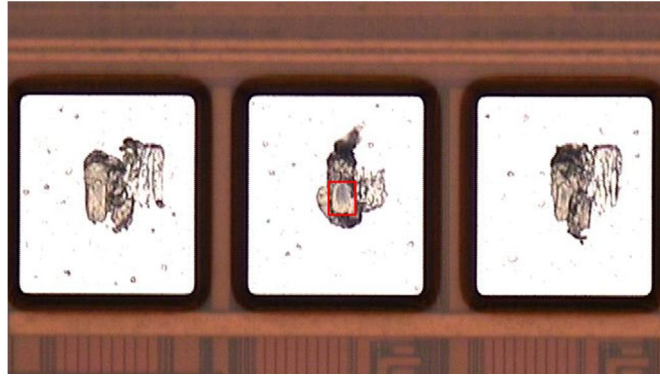


圖 10：針壓過大導致底部露出

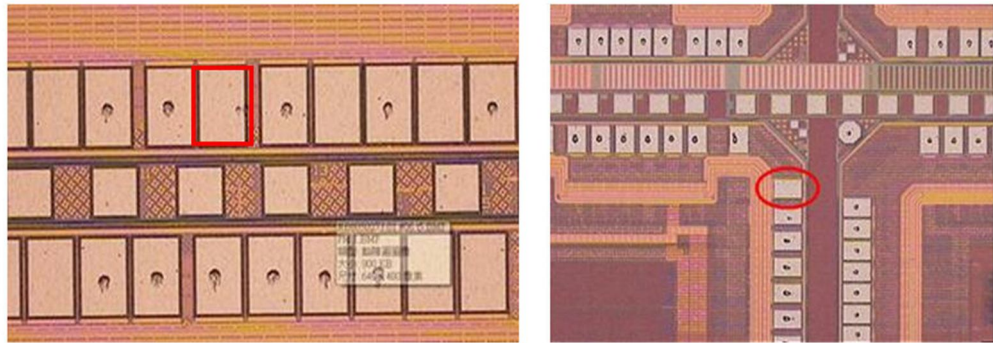


圖 11：多並測數探針未確實接觸的痕跡

無論是測試機台、針測台、針測卡等硬體設備的問題，或是人員的操作不當，未能確實將晶圓及探針卡定位等問題，所造成晶粒的針測結果不正常，而未能檢測出正確的良率，都歸類為誤宰 (Overkill) 現象。而常見造成晶圓測試誤宰的原因如下圖所示：

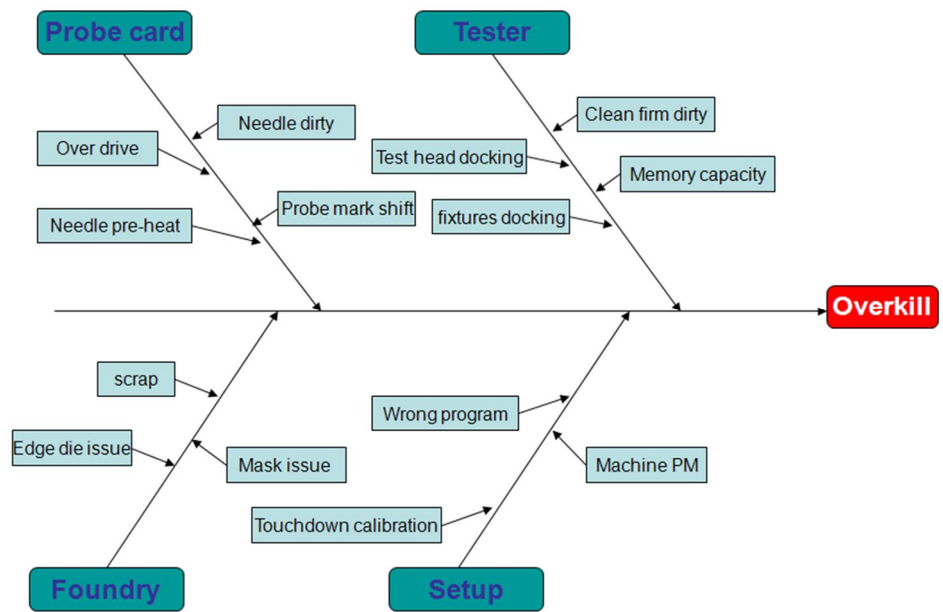


圖 12：晶圓誤宰原因魚骨圖

當誤宰狀況發生而導致測試良率低落時，必須仰賴具有經驗的工程師，對針測結果加以分析、進行實驗性的驗證、甚至換到不同機台上做交互比對，才能判斷是否為誤宰，以及後續是否要進行重測；此一過程不但需求高規人力，也需要花費冗長時間及成本。

一般情形下，工程師會先依據晶圓測試結果的幾何分佈 (wafer map) 以及針測摘要(Summary)作為初步判斷的依據，以下列出幾種常見的誤宰狀況所呈現出的針測結果特徵 (圖 13~16)：

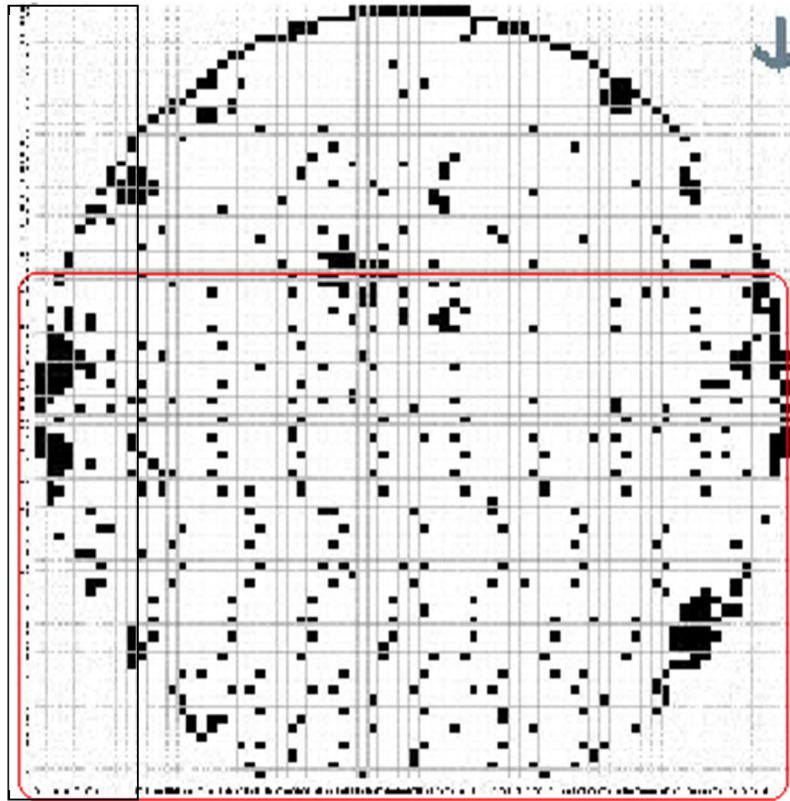


圖 13：異常的晶圓針測圖形

上圖的 wafer map 下半部，出現了呈現規則分佈的測試 fail。

S01A	S02A	S03A	S04A	S05A	S06A	S07A	S08A	S09A	S10A	S11A	S12A	S13A	S14A	S15A
286	253	296	291	284	290	281	282	286	280	295	88	289	284	283
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	3	7	3	1	3	1	1	5	3	2	5	3	1	1
0	43	1	0	1	2	0	0	0	11	0	0	1	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	12	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	26	34	40	37	18	24	22	20	26	34	208	26	20	19

圖 14：單一測試 site 的良率過低

而從圖 14 的針測摘要中顯示出第 12 號探針的良率明顯低於其他探針，表示可能在測試過程中，該探針沾粘了從晶圓上刮擦下來的殘屑，而未能在自動清針 (auto clean) 的程序去除，造成後續的晶粒針測結果失效。

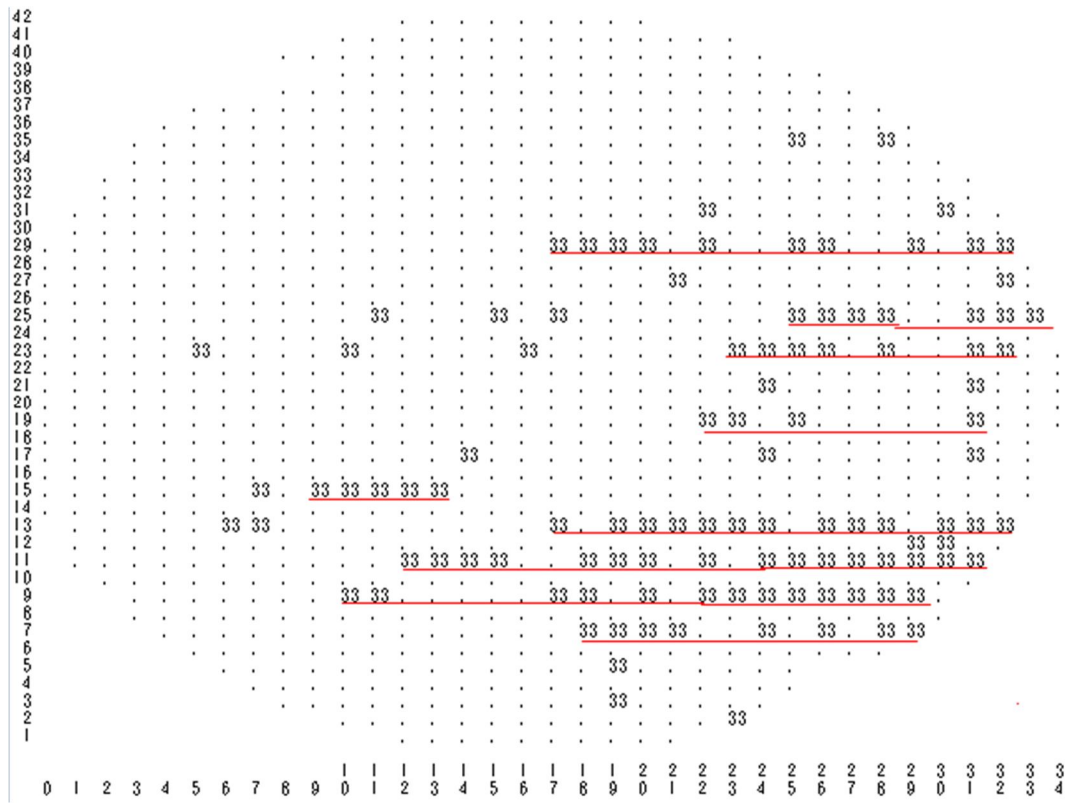


圖 15：與測試移動路徑類似的連續 defect

圖 15 的紅線部份為連續性測試 fail，並且出現的位置與測試路徑(針測卡的移動路徑)一致；此現象有可能是測試上的問題，也有可能是產品製程問題，通常必須針對該片晶圓進行點測 (Sampling Test) 分析，若該貨批有多片出現類似情形，則須挑選情況最嚴重的一片，換到不同機台並使用不同的針測卡，進行重測以驗證問題的所在。

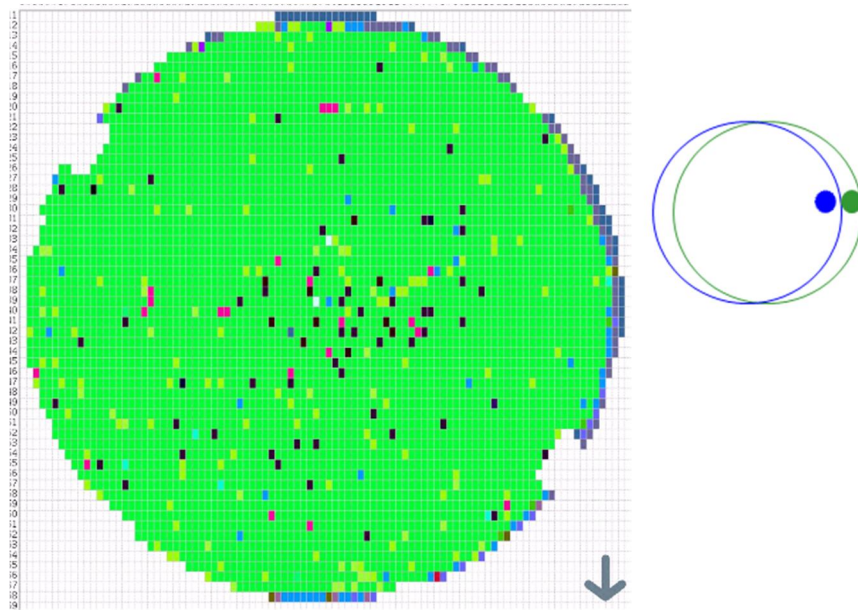


圖 16：測試 map 平移(人員設定問題)

圖 16 的針測結果右邊外圍出現大量的測試 fail，且 90% 以上均為 Open/Short/Leakage 的 fail，此一情形通常是因為操作人員對晶圓的 X、Y 平面定位未確實，造成整個 map 有平移的現象。

2.3 資料探勘

資料探勘(Data Mining)，也有人譯為資料採礦、數據挖掘；是近年來在資料庫應用領域的發燒話題。Frawley 等人認為資料探勘為資料庫中知識發掘 (Knowledge Discovery in Databases, KDD) 過程的一個核心步驟[6]；一般是指運用電腦科技，從大量資料中找出隱藏的訊息，例如趨勢(Trend)、特徵(Pattern)、相關性 (Relationship) 等等，然後轉變成對企業或組織在管理上具有意義的知識。Berry 及 Linoff 對它做了一個定義：「資料探勘是為要發現出有意義的樣型 (Pattern)或規則(Rule)，而必須從大量資料中，以自動或是半自動的方法來探索 (Explore)和分析資料的過程」[4]。另外，Kleissner 則解釋為：「資料探勘是一種新的且不斷循環的決策支援分析過程，它能夠從組合在一起的資料中，發現出隱藏價值的知識，以提供給企業專業人員參考」[8]。

資料探勘與傳統的統計分析有所不同：統計分析以假設及驗證為基礎，針對資料之間的關連性或統計學上不同標的加以分析，需要由具專業統計背景的專家針對統計結果來檢測，以及解釋其意義；而資料探勘則以「發現」為目的，著重於「模式辨認」，主要目的要找出資料中所「隱含」的具體規則，然後將結果交由可能不具專業統計知識的使用者(通常是高層決策人員，例如經理、總理等高階主管)，據以制定其因應決策。

通常將資料探勘的應用技術分為下列五種：

分類(Classification)

依據分析對象的屬性或特徵加以分門別類；先將類別群集的特徵清楚定義，然後透過訓練資料建立出模型，依該模型將未歸類的原始資料依其屬性或特徵加以分類。

分類的技術相當多，包含有決策樹(Decision Tree)、倒傳遞神經網路(Back-Propagation Neural network)、貝氏網路(Bayesian Network)、k 最近鄰分類(K-Nearest Neighbor Classifiers)、遺傳演算法(Genetic Algorithms)、模糊邏輯技術(Fuzzy Logic)等等。

推估(Estimation)

推估主要用於處理連續性的數值資料，以獲得某一屬性未知之值；例如依照信用卡持有人的薪資收入、年齡等資料，來推估其信用卡消費量。使用的技術包含統計方法上的相關分析、迴歸分析及類神經網路等等。

預測(Prediction)

預測是利用過去資料的觀察值，來預測未來的數值以及趨勢。主要使用迴歸分析、時間數列分析及類神經網路等等。

群集化(Grouping)

將一群異質的資料中，區隔出較有高度同質性的群組或子集合；也就是將群集之間的差異辨識出來，並挑選合乎個別群集的相似樣本。它與分類技術不同的是，群集化並沒有事先定義好的分群規則，而是由資料自身的相似性進行分組。

常用的演算法技術包括：質心法(K-means)、EM(Expectation-maximization Algorithm)及自我組織地圖法(SOM, Self-Organization Map)等。

關聯(Association)

從資料中發掘出那些事件總是同時發生，最典型的案例就是去分析在顧客在超級市場的同一次購物活動，對於其所購買的商品組成的相關性進行分析，因此有人稱關聯分析為購物籃分析(market-basket analysis)。

2.3.1 資料分類

資料探勘的分類技術，是利用一連串的輸入變數來對歷史資料分析，找出其特徵及規則，再對新的資料進行預測「分類」；例如從一個人的資產狀況、信用卡繳款記錄、貸款情形，來將他歸類到新銷售業務的潛在客戶等級。

Naïve Bayes 分類法

貝氏分類(Naïve Bayes)是根據貝氏定理(Bayesian Theorem) [3, 7]為基礎，以機率理論的概念計算資料是否屬於某一類別的可能性，屬於監督式學習的分類法。它先由訓練資料中計算各個類別與屬性的關係，並以此機率作為分類的依據，然後在測試資料時，對各類別都計算出一個機率值，並以機率最大者作為資料的歸屬類別。在某些領域的應用上，其分類效果優於類神經網路和決策樹。

貝氏定理是描述關於隨機事件 A 和 B 的條件機率以及邊緣機率的一則定理：

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

其中 $P(A|B)$ 是 B 發生的情形下，A 事件發生的機率。

現在假設 W 代表某筆資料的一組特徵，則具有特徵 W 的資料被分類到類別 C_j 的機率定義：

$$P(C_j|W) = \frac{P(W|C_j)P(C_j)}{P(W)}$$

$P(W)$ 表示特徵 W 的機率， $P(C_j)$ 表示分類在 C_j 的機率， $P(W|C_j)$ 則表示類別 C_j 中具有特徵 W 的機率。在此採用獨立假設，假設每個特徵都是獨立的，且對於每個類別 C_j 而言， $P(W)$ 都是相同的，則上式可改寫為：

$$P(C_j|W) = P(W_1|C_j) \times P(W_2|C_j) \times P(W_3|C_j) \times \dots \times P(W_n|C_j)$$

計算出機率最大的 $P(C_j|W)$ 就是最有可能的類別。

支持向量機(Support Vector Machines)

支持向量機(Support Vector Machines, SVM) [5]是由 Vapnik 及 Cortes 在 1979 年提出的一種以統計學習理論(Statistical Learning Theory) [14]為基礎的機器學習演算法，它屬於監督式學習。它主要的特色在於將變數對應關係轉到高維度空間，並在該空間內找出一個最佳化的分隔超平面(Optimal Separating Hyper-plane, OSH)，這個超平面可以讓被分隔的兩類別之間的邊緣(margin)距離最大化，因此 SVM 也被稱為最大邊緣分類器。

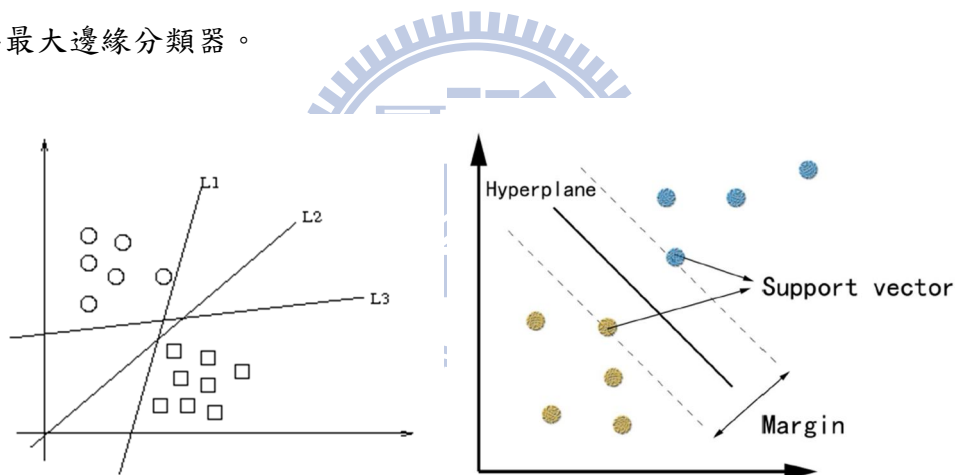


圖 17：多個超平面可切開類別，但只有一個能達到最大分割

SVM 演算法超平面的數學形式可以寫成：

$$\mathbf{w} \cdot \mathbf{x} - \mathbf{b} = 0$$

其中 \mathbf{x} 是在超平面上的點，而 \mathbf{w} 是垂直於超平面的向量，而 \mathbf{b} 則表示位移值；SVM 的計算過程中，會先將訓練資料透過上述函數計算出一條線，若線上的點帶入此函數所計算得到的值為 +1 或 -1，則位於該線上的點為支持向量(Support Vector)。由於計算之後可能得到多個超平面或支持向量將資料一分為二，若兩條

虛線間的邊界 (margin)越大，越能有效率的切割資料集，因此 SVM 的最終目標就是最出具有最大邊界的超平面。

決策樹(Decision Tree)推論法

決策樹(Decision Tree)是在人工智慧以及機器學習領域中很受歡迎的一種預測方法，屬於監督式學習法；透過事先選定的目標函數，依據此目標函數，將資料進行分支，以形成樹狀結構；任何的一個內部節點(node)都是一個測試的屬性，而每一條分支(branch)都代表了測試的結果，最後的葉節點(leaf node)則代表各類別的分佈情形。

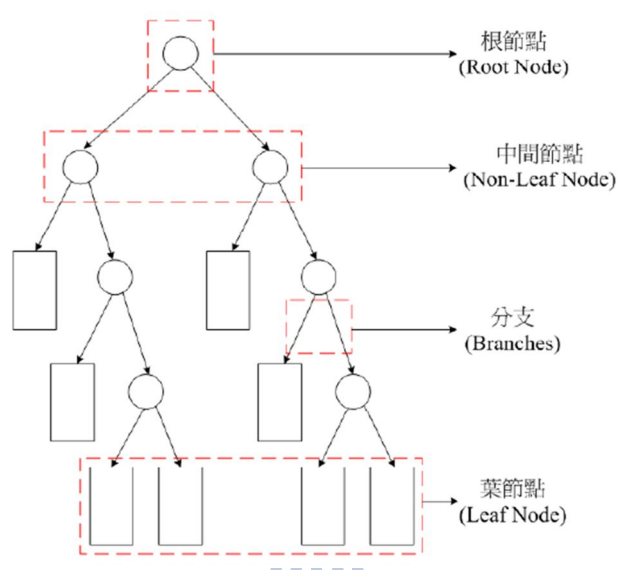


圖 18：決策樹基本架構

決策樹使用的 ID3(Iterative Dichotomiser)演算法[10]，由 John Ross Quinlan 在 1986 年所提出，它廣泛用於處理離散型的資料型態。ID3 演算法最重要的就是能降低決策樹的複雜度並兼顧分類的正確性。

在建構決策樹過程中，ID3 以資訊獲利(Information Gain)為依據，選擇最佳的屬性當成節點，使得選擇分類屬性的準則偏向選擇屬性值較多的屬性，但該屬性在實際決策上，可能並不具決策價值。

為了改善此缺失，Quinlan 提出 C4.5 [11, 12] 的決策樹歸納法：其主要分為兩部份，一為分類的標準，依據獲得量比值(Gain Ratio)來計算；另一為修剪的標準，依據錯誤率為基礎之修剪(Error Based Pruning)。

C4.5 與 ID3 最大的差別是其屬性離散化的動作是在各節點內動態決定的，亦即在不同節點中，特徵值離散化的結果不盡相同。

C4.5 對於連續性數值屬性節點分割方法，是先將物件集合依此屬性作排序，再依次找出鄉鄰二個物件的屬性終點，成為分割點 (Cut point)，然後以評估函數計算，能得到最佳值者即依此屬性中點作二元分割。

倒傳遞神經網路(Back-Propagation Neural Network)

類神經網路是仿造人類大腦組織及其運作方式的一種分類技術，神經網路是一種運算模型，由大量的節點(或稱「神經元」，或「單元」)和之間的相互聯接構成。通過對局部情況的對照比較(而這些比較是基於不同情況下的自動學習和要實際解決問題的複雜性所決定的)，它能夠推理產生一個可以自動識別的系統。

它必須先經過訓練(training)的程序，進行反覆學習，直到每個輸入值都能正確對應到所需的輸出結果；訓練的目的，就是讓類神經網路的輸出值能盡量的接近目標值。因此，在進行類神經網路的學習之前，必須先準備一個訓練樣本(Training Pattern)，以做為參考用。

目前類神經網路模式有數十種，在實際應用上以倒傳遞網路(Back-Propagation Network, BPN) 最為普遍，由 Werbos 首先在 1974 年的博士論文中，提出在感知機加入隱藏層的學習演算法[15]，但直到 1985 年 Parker 再次提出倒傳遞神經網路[9]，隔年由 Rumelhart 等人[13]再發表一篇文章後才廣為人知。它使用多層感知器 (Multi-Layer Perception) 排列方式，形成一個類神經網路架構，包含輸入層、隱藏層以及輸出層，每一層包含若干個處理單元；輸入層用以輸入外來的訊息，而輸出層則將計算過後的數學函數值輸出，其間可能包含若干層隱藏層，表示類神經網路處理單元間的交互作用。

2.3.2 關聯法則

關聯法則用來分析在資料庫中不同變數或個體之間的關係程度(機率大小)，最早在 1993 年由 Agrawal 等人[1]所提出，他對關聯規則的定義如下：

關聯規則中有兩個重要的參數，分別為支持度(Support)和信賴度(Confidence)，用以判斷某關聯規則是否有意義或值得探討。

支持度表示符合某關聯規則的交易次數比例；對關聯規則 $A \rightarrow B$ (購買 A 商品後也同時購買 B 商品)來說，此規則在 D 中的支持度為「有購買 A 商品及 B 商品的交易佔總交易記錄百分比」，即機率 $P(A \cup B)$ 。

信賴度是指包含 A 的所有記錄交易中，同時包含 B 的百分比，即條件機率 $P(B | A)$ 。

一般以此方法計算出來的規則數非常多，因此必須預先設定支持度和信賴度的門檻值，同時符合該門檻值的規則才會被挑出來。

實作上以Agrawal 與Srikant 所提出的Apriori演算法[2]最具代表性，它利用循序漸進的方式，找出資料庫中項目的關係，以形成規則。其優點為結構簡單、易於理解，沒有複雜的推導，例缺點為計算項目過多，導致執行效能緩慢。

關聯規則最著名的應用範例為「購物籃分析」(Market Basket Analysis)，是美國沃爾瑪百貨(WAL-MART)的管理人員分析銷售數據時發現，跟尿布一起購買最多的商品竟然是啤酒，後來透過市場調查才得知，原來美國的太太常叮嚀丈夫在下班前幫嬰兒買尿布，而 40% 的先生在買完尿布後，又會隨手拎兩罐啤酒。既然尿布和啤酒一起購買的情形特別多，威名百貨乾脆就在賣場把啤酒和尿布擺在一起，結果尿布和啤酒的銷售量雙雙增加三成！

購物籃分析後來成為零售業經常應用的一種分析技術，但不僅限於零售業，即使用在金融、保險及電信業等等，也證明是有效的方法。

第 3 章 研究架構設計

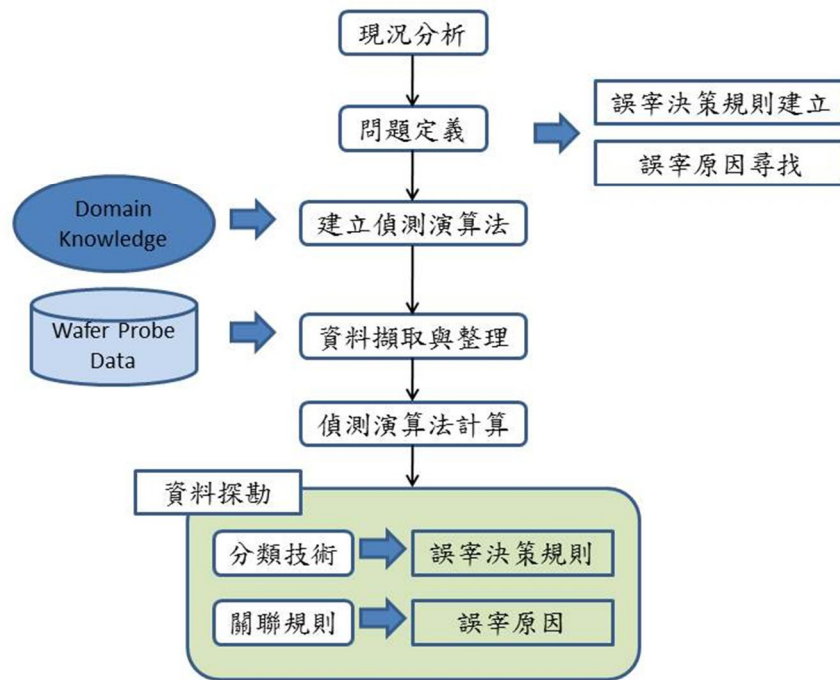


圖 19：系統資料探勘架構

本研究所設計之研究架構如圖 19 所示，乃經過業界現況分析、問題定義、與專家討論建立誤宰偵測演算法、資料擷取整理以及使用偵測演算法進行計算；然後再將取得的計算結果，透過資料探勘的分類技術進行分析，確認偵測演算法的準確性；然後再以關聯規則分析技術，試圖找出與誤宰現象有相關的環境以及設定參數。

晶圓針測資料的來源，以業者實際生產線的資料庫為主，包括第一次針測及重測後的良率及差異值、測試站別、所使用的機台、產品名稱、操作人員等等；再與領域專家討論設計出和誤宰特徵相關的演算法，然後針對前述針測資料套用這些演算法，並將演算法所計算出的結果，合併成為後續資料探勘分析所使用的資料。

在經過資料探勘技術分析之後，選出最適合的方法，並將所推論出的決策規則，與業界專家進行討論，將此決策規則實際轉化為晶圓測試過程中有意義的知識。

3.1 現況分析

對晶圓測試業者來說，在整個測試過程中，是否能正確的反應出晶圓該有的良率，是衡量業者技術能力及品質的重要指標；然而，由於進行晶圓測試時，所使用的測試機台、探針機、探針卡等硬體設備的穩定性，以及人員操作的正確性，都有可能造成誤將良品判斷為壞品的誤宰現象出現。

個案公司已具備測試機台資料自動化傳送的機制，當測試機台完成一片晶圓的測試程序後，便會將相關的資料透過網路上傳到業者的 TPAS(Testing Parameter Analysis System)系統，經過處理後，將資料儲存到資料庫中，簡易流程如圖 20 所示：

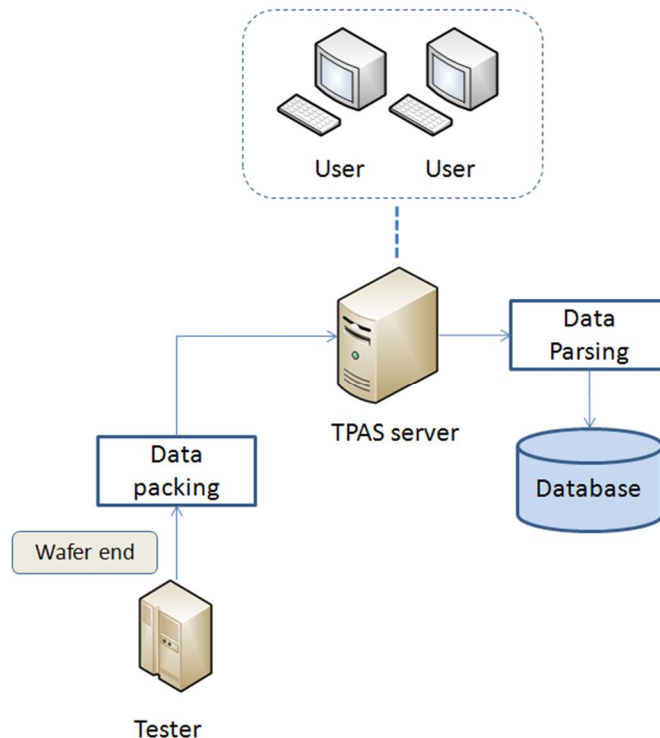


圖 20：測試資料傳輸架構圖

在上圖中的 Data parsing 程序中，現有系統便已針對每次的晶圓測試資料，整理出 Bin map、針測結果摘要、良率、探針卡各 site 的良率(針對多個並測數的探針卡)...等等資訊，並儲存在資料庫中，以便使用者查詢以及將資料轉換成客戶所要求的各式報表。

業者在其系統內針對各產品制訂了良率標準範圍，當晶圓針測良率超出該標準範圍時，便會在系統上出現警示，此時生產線人員便會通知值班工程師，針對該針測結果進行檢視，分析是什麼原因造成良率異常。

然而，對於因為誤宰現象所引起的良率異常，仍必須仰賴資深工程師的經驗進行判定；經過與業界工程師實際訪談後，發現晶圓誤宰對不同產品所呈現的資料特徵也有所不同，對經驗較不足的工程師而言，不易判定是否為誤宰，而必須再進行實驗性的點測驗證、或交換不同硬體組合重測後，才能依重測後的良率判斷是否為誤宰現象。

若能針對誤宰現象所常呈現出的資料特徵，設定出對應的演算法，然後在 Data parsing 程序之後進行運算，並將計算結果也一併儲存到資料庫內，之後便可以依據這些演算法的計算結果進行分析，例如以決策樹方法建立起與誤宰的關係規則，日後便可當成判斷的輔助資訊。

3.2 問題定義

對於業者來說，晶圓測試時發生誤宰現象，不僅代表者重測所必需耗費的成本及時間的浪費，更重要的是影響到客戶眼中，業者的技術能力是否穩定的形象。而在判斷是否為誤宰現象方面，也須仰賴有經驗的工程師，進行實驗性的分析，對高規人力的需求及養成，也是一個不易達成的議題。

經過與業者的領域專家訪談之後，可歸納出對業者最迫切需要解決有兩個問題：

1. 誤宰現象的判斷(Trouble Shooting)

當異常良率出現時，如何快速而準確的判斷是否為誤宰？

2. 造成誤宰的原因分析(Root Cause Finding)

能否找出造成誤宰現象的原因？例如：機台與探針卡的匹配問題、人員操作不當...

3.3 誤宰偵測演算法的建立

經過與領域專家的討論，針對發生誤宰的晶圓針測資料，常會出現某些特徵，可當成工程師判斷的參考依據。例如圖 21 所出現的 Bin Map，在 map 圖的下半部，出現特定的 fail 排列方式：

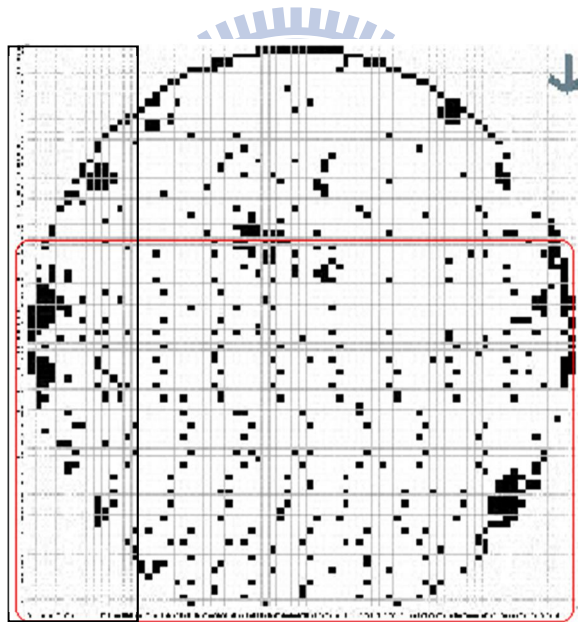


圖 21：特定排列方式的 bin map

再觀察其針測摘要（圖 22），可以發現探針卡上的第 12 號針的良率特別低 (fail 顆數特別高)：

3.4 資料萃取

本研究的資料取自業者於 2011/10~2011/12 間的針測資料，其中包含第一次針測及重測後的良率及差異值、測試站別、所使用的機台、產品名稱、操作人員等等，再對每一片晶圓的針測資料套用上述誤宰偵測演算法，並將計算結果也儲存起來，成為後續資料探勘時使用的資料集。

依據業者的定義，重測後若能救回超過 2%的良率，便可認為是誤宰，因此在資料集中加入一個 'Overkill' 的欄位來標示，此一欄位的值，便是由重測後的良率來計算取得：

wafer_id	yield	yield_delta	overkill	tester	prober	probe_card	test_pg	claim_user_id	dut_cnt	temperature	v1	v2	v3	v4
SSLY907	97.6	0.3	N	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	-	-
SSLY908	97.2	0.6	N	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	-	-
SSMR301	94.8	3	Y	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.4	-	-	Y	-
SSMR302	93.3	3.9	Y	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.2	-	-	Y	-
SSLY909	96.2	1.4	N	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	-	-
SSMR303	93.1	3.1	Y	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.1	-	-	Y	-
SSLY910	94.3	1.9	N	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	Y	-
SSMR304	88.9	5.4	Y	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.1	-	-	Y	-
SSMG907	95.3	0.7	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A2300	32	25.2	-	-	-	-
SSMG908	93.7	1.5	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A2300	32	25.2	-	-	Y	-
SSMG911	95.1	1	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A2300	32	25.3	-	-	-	-
SSLY911	91.9	2.9	Y	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	Y	-
SSMG912	93	1.8	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.6	-	-	Y	-
SSMG913	92.6	3.9	Y	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.5	-	-	Y	-
SSCH116	94.4	3.9	Y	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	32	25.3	-	-	Y	-
SSMG914	94.4	1.2	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.5	-	-	Y	-
SSMG915	92.9	1.4	N	TDH26A	TSK083	4582/1/28	PhoenixC2Up9_RevBx32_r6.xls	A3033	32	25.6	-	-	Y	-
SSLY912	96.9	0.6	N	TDH24A	TSK084	4582/1/26	PhoenixC2Up9_RevBx32_r7_eng1_FRDY.xls	A2719	32	25.1	-	-	-	-
SSCH117	90.9	6.9	Y	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	32	25	-	-	Y	-
SSMR301	97.1	0.7	N	TDH26A	TSK083	4582/1/25	PhoenixC2Up9_RevBx32_r6.xls	A1748	32	25.8	-	-	-	-
SSMR302	96.3	0.9	N	TDH26A	TSK083	4582/1/25	PhoenixC2Up9_RevBx32_r6.xls	A1748	32	25.4	-	-	-	-
SSMR303	94	2.2	Y	TDH26A	TSK083	4582/1/25	PhoenixC2Up9_RevBx32_r6.xls	A1748	32	25.2	-	-	Y	-

救回良率

偵測演算法計算結果

圖 23：資料處理範例

表 1：資料欄位說明

欄位名稱	說明
wafer_id	晶圓編號
yield	晶圓在當次測試的良率
yield_delta	重測與重測前的良率差
overkill	是否為誤宰
tester	測試機台
prober	針測台

probe_card	針測卡
test_pg	測試程式名稱
claim_user_id	操作人員工號
dut_cnt	針測卡並測數
temperaure	測試時溫度
V _n	偵測演算法編號及計算結果

3.5 資料探勘

本研究使用 Weka [16] 套件軟體來進行資料探勘分析的實作部份。Weka (Waikato Environment for Knowledge Analysis) 是由紐西蘭 Waikato 大學所發展的機器學習軟體套件，它包含了許多目前常見的機器學習演算法；Weka 軟體可對資料做預處理及轉換，並提供分類、迴歸、群集、關係法則等等的分析演算法，並可以視覺化的方法提供分析結果。

本研究使用 Weka 版本為 3.6.6，相關資料探勘技術與對應的 Weka 演算法如表 1：

表 2：資料探勘技術與 Weka 演算法對照表

資料探勘技術	Weka 演算法
Naïve Bayes	weka.classifiers.bayes.NaiveBayes
C4.5 Decision Tree	weka.classifiers.trees.J48
BPN	weka.classifiers.functions.MultilayerPerceptron
SVM(SMO)	weka.classifiers.functions.SMO
Association rules	weka.associate.Apriori

第 4 章 實驗分析

本研究的目標個案公司為一專業晶圓測試代工廠，其營業收入幾乎全為晶圓測試所得；測試誤宰對該公司而言，不僅造成成本及生產時間的耗費，也是影響公司對外形象的重要因素。

由於高科技產品的競爭越趨激烈，產品的生命週期越來越短，任何製造生產過程中的延誤，都有可能造成商業利益的損失；當晶圓測試過程出現誤宰現象時，工程師驗證及判斷所花費的時間，便會成為生產延誤的累積變數。因此，首要任務就是要能協助工程師對誤宰現象的判斷。

4.1 實驗一：針測誤宰決策規則建立

4.1.1 資料準備

首先，我們擷取個案公司的 2 個產品型號，在 2011/10~2011/12 期間，曾經進行過重測流程的晶圓資料，並且取得其重測後與重測前的良率差異值；以業者的內部管理定義，當重測後所救回來的良率，若是超過百分之二，便可視為是誤宰。因此，我們在原始資料中，加入一個衍生的欄位-Overkill，其值便由救回的良率差來呈現。

wafer_id	yield	yield_delta	overkill	tester	prober	probe_card	claim_user_id	dut_cnt
SSAY317	98.3	0	N	TDH31A	TSK081	4582/1/27	C0427	32
SSAY318	98.2	0	N	TDH31A	TSK081	4582/1/27	C0427	32
SSAY301	97.2	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY301	95.7	1.6	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY302	97.6	0	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY304	98.6	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY304	96.5	2	Y	TDH31A	TSK081	4582/1/23	C0331	32
SSAY305	98.3	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY305	96.5	1.9	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY307	98.1	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY307	95.1	2.9	Y	TDH31A	TSK081	4582/1/23	C0331	32
SSAY308	98.3	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY308	96.5	1.9	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY309	98.2	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY309	87.4	10.8	Y	TDH31A	TSK081	4582/1/23	C0331	32
SSAY310	98.4	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY310	93	5.3	Y	TDH31A	TSK081	4582/1/23	C0331	32
SSAY311	97.9	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY311	96	1.9	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY312	97.8	0	N	TDH31A	TSK081	4582/1/24	C0384	32
SSAY312	94.7	3.1	Y	TDH31A	TSK081	4582/1/23	C0331	32
SSAY313	97.6	0	N	TDH31A	TSK081	4582/1/23	C0331	32
SSAY314	98.5	0	N	TDH31A	TSK081	4582/1/24	C0384	32

圖 24：額外增加的 Overkill 欄位

4.1.2 資料處理

上述取得的晶圓針測資料，套用上針對誤宰資料特徵所設計的演算法，並將各個偵測演算法所計算出的結果記錄下來，回存到該資料集中。在此我們以”v+數字”代表各個不同的誤宰偵測演算法，計算結果若是 ‘Y’，表示該筆資料觸發了該項偵測演算法。

overkill	tester	prober	probe_card	claim_user_id	dut_cnt	temperature	v1	v2	v3	v4	v5	v6	v7	v8
N	TDH31A	TSK081	4582/1/27	C0427	32	25.6	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/27	C0427	32	25.3	-	-	-	-	-	-	Y	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/23	C0331	32	89.8	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/23	C0331	32	90	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/24	C0384	32	90.1	-	-	-	-	-	-	-	Y
Y	TDH31A	TSK081	4582/1/23	C0331	32	90.1	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90.1	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/23	C0331	32	90.1	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90.1	-	-	-	-	-	-	-	Y
Y	TDH31A	TSK081	4582/1/23	C0331	32	90.1	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90.1	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/23	C0331	32	90.1	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90	-	-	-	-	-	-	-	Y
Y	TDH31A	TSK081	4582/1/23	C0331	32	90.1	-	-	Y	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90	-	-	-	-	-	-	-	Y
Y	TDH31A	TSK081	4582/1/23	C0331	32	90.2	-	-	Y	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	90	-	-	-	-	-	-	-	Y
N	TDH31A	TSK081	4582/1/23	C0331	32	90.2	-	-	-	-	-	-	-	-
N	TDH31A	TSK081	4582/1/24	C0384	32	89.9	-	-	-	-	-	-	-	Y

圖 25：誤宰偵測演算法的計算結果

4.1.3 資料探勘

取得上述預先整理過後的資料，本研究使用 Weka 這套 OpenSource 的軟體，將前面步驟準備好的資料，以 CSV (Comma Separated Values) 檔案的格式，匯入 Weka 軟體進行分析。

我們採用讓訓練資料集與驗證資料集綜合方式進行，設定百分比為 70% 的訓練資料集，讓 Weka 先針對訓練資料集隨機抽樣，進行訓練模式推導，再以剩下 30% 的驗證資料集代入訓練模式，試算以偵測演算法來推論誤宰現象的準確度如何。各分類技術的參數設定如表 3 所示：

表 3：各演算法的參數設定

Weka 演算法	參數設定
Naïve Bayes	weka.classifiers.bayes.NaiveBayes
C4.5 Decision Tree	weka.classifiers.trees.J48 -C 0.25 -M 2
BPN	weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
SVM(SMO)	weka.classifiers.functions.SMO -C 1.0 0L 0.0010 -P 1.0E-12 -N 0 -V -1 -W -1 -K "weka.classifiers.functions.supportVector.PolyKernel-C 250007 -E 1.0"
Association rules	weka.associate.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S 1.0 -c -1

分類技術的執行環境為 Windows 7, AMD Athlon 64x2 5200+ (2.7GHz),

從各分類方法分析偵測演算法對不同產品針測結果的計算資料，可以看到其準確度最差為 SVM 的 81.75%，最高為 C4.5 及 BPN 的 93.20%。足見這些針測演算法對誤宰現象的預測，具有相當高的準確度。C4.5 決策樹及倒傳遞神經網路這兩種分類技術，所推論出的準確度非常相近，但 C4.5 決策樹只花了不到 1 秒，而倒傳遞神經網路卻各花了 5 分 48 秒與 3 分 49 秒才計算出來；對於想要應用於生產線系統的功能而言，反應時間是很重要的一個要素，決策樹在此佔了很大的優勢。

表 4：使用不同分類方法檢驗偵測演算法的準確度

分類方法 \ 產品	Naïve Bayes	Decision Tree	BPN	SVM
A (1369 筆)	83.45%	86.86%	87.10%	81.75%
B (884 筆)	89.43%	93.20%	93.20%	92.83%

對 A 產品所分析出來的決策樹規則如下圖所示，其中 v3、v5、v24 及 v31 便可分辨出大部份的誤宰：

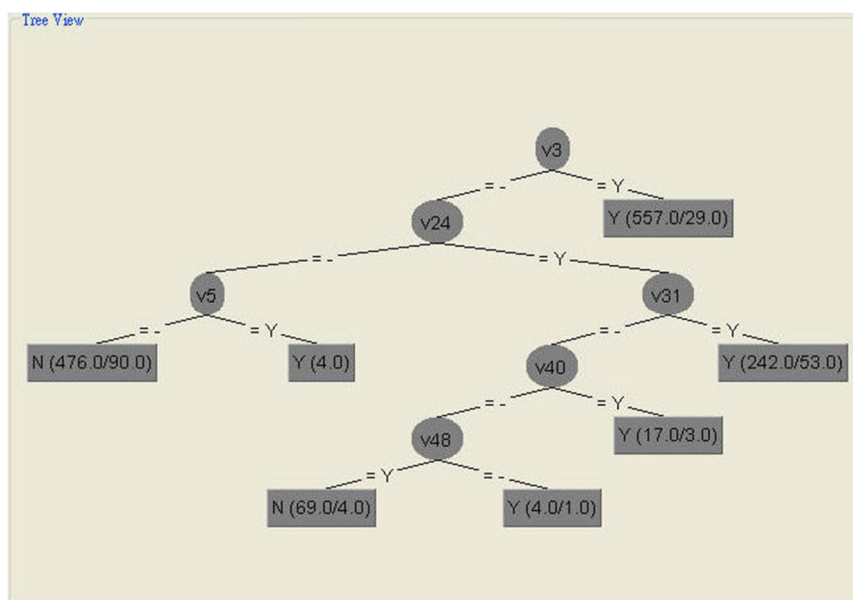


圖 26：A 產品決策樹規則

對 B 產品所分析出來的決策樹規則如下圖所示，可看出 v3、v5 以及 v24 便可分辨出大部份的誤宰記錄：

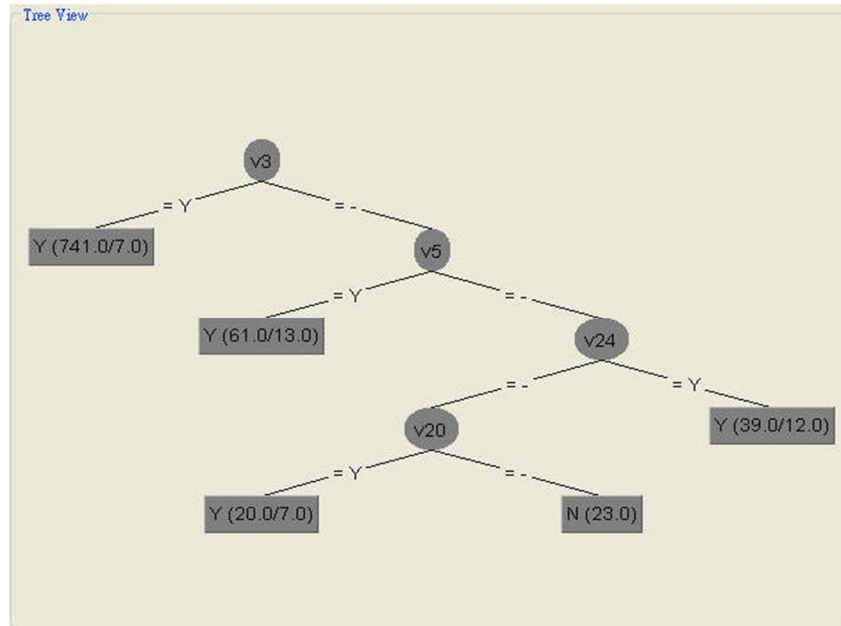


圖 27：B 產品決策樹規則

從上述對 A, B 兩產品所推論出的決策樹規則來看，v3、v5、v24、及 v31 這四種偵測演算法，就足以推斷出大部份有誤宰情形的針測資料，實務上可使用所推論出的決策樹規則來當成判斷依據：

V3：檢查良率是否低於某一標準值，此種偵測條件適用於製程已穩定成熟的產品，其良率也應該要很穩定，若出現低良率的情形，表示可能針測的過程有問題。

V5：此偵測條件會檢查由廠內工程師所指定的特定 Bin 編號，其失效晶粒顆數是否超過某特定比例，通常若測試程式是由廠內所撰寫，便能由工程師自行將該測試項目加入到程式中。

V24：此偵測條件須配合針測程式，在針測過程中，進行特殊的項目測試，若該特定項目的 fail 顆數超過設定值，則視為該偵測條件被觸發。

V31：單一 site 的良率與其他 site 的差異值過大(Site Low Yield)，此一狀況通常是因為針壓不足或探針上有異常殘屑存在所導致。

V40：主要是比對 CP1（晶圓第一道測試）與 CP2（晶圓第二道測試）的良品晶粒顆數差異，若超過設定的標準值，就觸發此一偵測條件。正常狀況下，CP1 若被判為良品的晶粒，在 CP2 階段的測試不應該被判斷為壞品。

4.2 實驗二：針測誤宰原因分析

在實驗二中，我們將針對晶圓針測誤宰的原因加以分析。

導致針測誤宰的原因有很多種，包括測試機台、針測卡、測試程式、人員的操作程序等，甚至於這些變數的組合：例如 A 測試機台與 B 針測卡的搭配組合等等。一般較常見的針測誤宰原因有：

測試機台：測試頭定位、配件定位、是否有定期校驗。

針測卡：針頭不潔、針壓不足或過度、針頭預加溫。

人員操作：是否按照 SOP 操作、試測前的對齊動作。

其他：前段晶圓製造刮傷、光罩對齊問題。

我們的目標是透過針測過程中的這些變數，試圖找出與誤宰狀況之間是否有關聯性；因此採用資料探勘技術中的關聯規則(Association Rule)來進行分析。

4.2.1 問題定義

從晶圓測試廠的針測資料中，我們可以取得測試機台、針測卡、測試程式、操作人員、測試溫度等等，本實驗將找出這些變數與針測誤宰現象間的關係，以作為業者在製程改善的依據。

4.2.2 資料準備

本實驗使用業者某一產品型號於 2011/11~2011/12 間共三個月的重測資料，資料變數包含測試機台編號、針測機編號、測試站別、針測卡編號、測試程式及操作人員等等。同樣的，依照重測後所救回來的良率，若是超過百分之二，便可視為是誤宰。因此，我們在原始資料中，加入一個衍生的欄位-Overkill，其值便由救回的良率差來呈現。

op_name	tester	prober	probe_card	test_pg	claim_user_id	dut_cnt	temperature	overkill
CP1	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	32	25.6	N
CP1	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	32	25.3	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	89.8	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90.1	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90.1	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90.1	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90.1	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.2	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.2	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	89.9	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90	Y
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	32	90.1	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	32	90.1	N

圖 28：資料萃取範例

表 5：誤宰原因分析用資料欄位

欄位名稱	說明
op_name	針測站別名稱
tester	測試機台編號
prober	針測台
probe_card	針測卡
test_pg	測試程式
claim_user_id	操作人員工號
dut_cnt	針測卡並測數
temperaure	測試時溫度
overkill	衍生欄位（是否為誤宰）

由於 Weka 的 Apriori 演算法只能針對“類別型”資料進行分析，因此我們將 temperature 此欄位，依據該產品的測試參數規定，轉換成類別型資料：

CP1 :

temperature <= 22 , → 'belowTempSpec'

22 < temperature < 28 , → 'normal'

temperature >= 28 , → 'aboveTempSpec'

CP2 :

temperature <= 88 , → 'belowTempSpec'

88 < temperature < 92 , → 'normal'

temperature >= 92 , → 'aboveTempSpec'

由於轉換後的資料均為 'normal'，表示機台進行針測時的溫度都在該產品的規範之內，因此在我們這次的實驗中可從分析參數剔除。

另一欄位 dut_cnt 表示針測卡的並測數，由於此一產品只使用 32 並測數的針測卡，所以也可以從此次實驗的分析參數剔除。

op_name	tester	prober	probe_card	test_pg	claim_user_id	overkill
CP1	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	N
CP1	TDH31A	TSK081	4582/1/27	PhoenixC2Up9_RevBx32_r6.xls	C0427	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	Y
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N
CP2	TDH31A	TSK081	4582/1/23	PhoenixC2Up9_RevBx32_r6.xls	C0331	N
CP2	TDH31A	TSK081	4582/1/24	PhoenixC2Up9_RevBx32_r6.xls	C0384	N

圖 29：準備進行關聯規則分析的資料欄位

4.2.3 資料探勘

取得上述整理過後的資料，對這些資料進行關聯規則的分析；我們選擇 Weka 軟體的 Apriori (weka.associations.Apriori) 演算法來進行分析。設定參數如下：

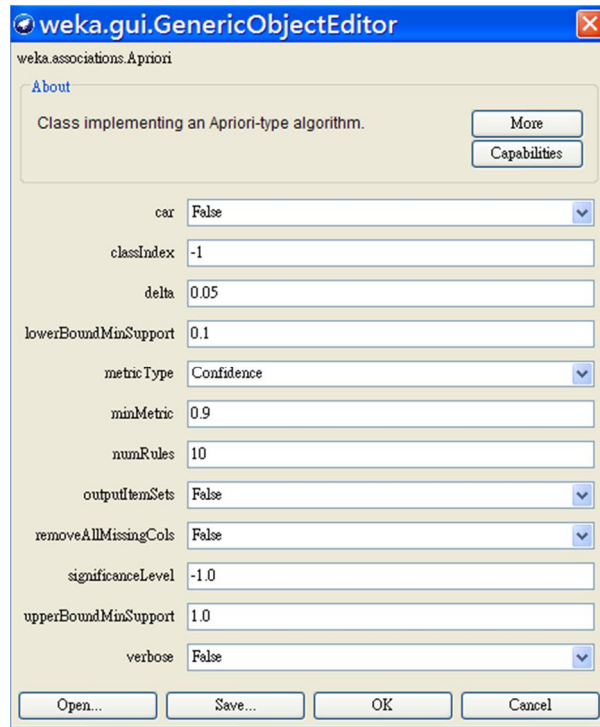


圖 30：Apriori 參數設定

lowerBoundMinSupport (最小支持度)：預設值為 0.1

metricType：'Confidence'

minMetric (最小信賴度)：0.9

表示要從資料找出符合信賴度 90% 以上且支持度為 10% 到 100% 之間的規則。

執行結果如下：


```

Minimum support : 0.1 (139 instances)
Minimum metric <confidence> : 0.9
Number of cycles performed : 18

Generated sets of large itemsets :
Size of set of large itemsets L(1) : 16
Size of set of large itemsets L(2) : 34
Size of set of large itemsets L(3) : 30
Size of set of large itemsets L(4) : 7

Best rules found :
1. tester=TDH26A, prober=TSK083 259 ==>overkill=Y 259...conf : (1)
2. prober=TSK083 overkill=Y 259 ==>tester=TDH26A 259...conf : (1)
3.tester=TDH31A 155 ==> test_pg=PhoenixC2Up9_RevBx32_r6.xls 155...conf :
(1)
4.test_pg=PhoenixC2Up9_RevBx32_r7.xls,overkill=N 151 ==>op_name=CP1
151... conf : (1)
5. tester=TDH26A, test_pg=PhoenixC2Up9_RevBx32_r7.xls 161 ==>overkill=Y
156... conf : (0.96)

```

圖 31：關聯規則分析結果

我們可以看到從上述的分析所推論出的規則中，去除與 overkill 無關的規則 3 後，剩下的 1,2,5 三條規則中，可以看出 'tester=TDH26A', 'prober=TSK083'，與 overkill 現象存在著關聯性。而在個案公司的機台配置中，測試機台與搭配的針測台幾乎不做更換，也就是可以視為一體。

將上述分析的結果反應給業者，業者先在廠內進行交叉驗證，將 TDH26A 所搭配的針測機 TSK083 與其他針測機交換，進行試測數片後，發現仍會出現誤宰的狀況，更加確定問題是出在 TDH26A 機台上；後於 2012 年 1 月份連絡設備廠商進行檢修後，檢出 TDH26A 上的其中一個測試 Channel 電路板可能是主因，該電路板上有二顆鋁質電容出現膨脹現象，在遇到較大電流的測試項目時，常會出

現不正常的結果，但在一般的測試項目中，則不會出現問題，因此難以掌握；經過檢修更換電路板之後，再擷取 TDH26A 機台後續 3 個月的針測記錄，可以看出誤宰的狀況已大幅改善；證明經由關聯規則所推論出的結果，對於業者在製程管理的改善方面，可以提供有效的輔助資訊。

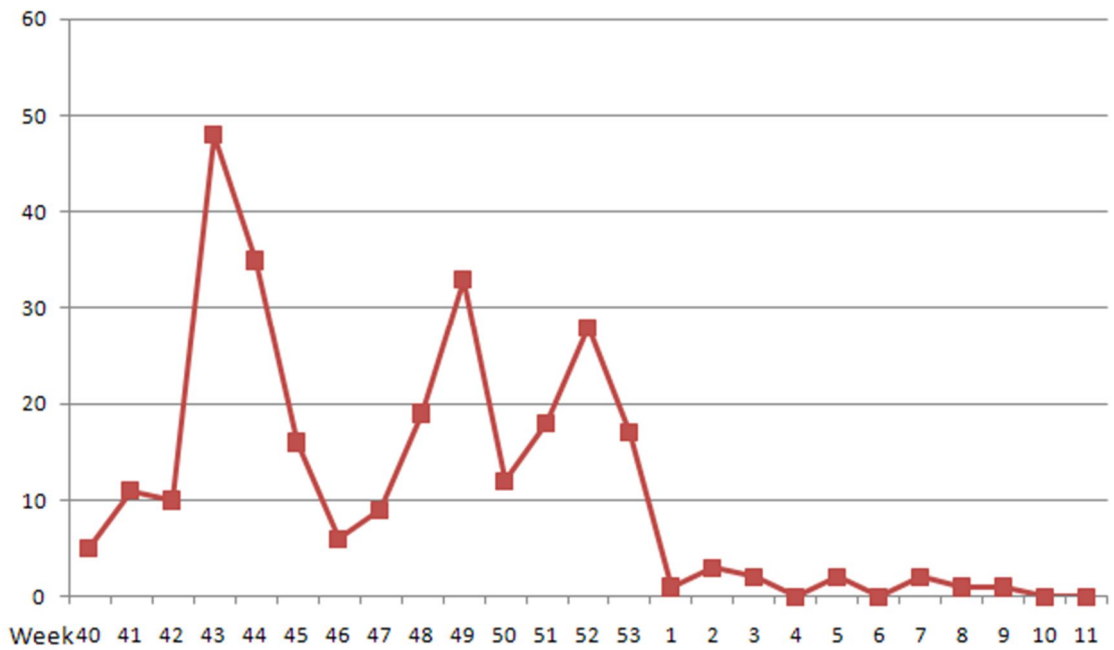


圖 32：問題機台每週誤宰筆數統計圖

第 5 章 結論與未來研究方向

台灣半導體產業經過多年專業分工後，雖已形成上下游的群聚，然而面對競爭全球化的趨勢，更必須建立起特有的競爭力；減少針測誤宰現象的出現，乃是證明測試能力的最佳方式。尤其在所有業者都競相提高軟硬體設備的投資時，若能擁有在管理及資料分析方面的「軟實力」，便能在減少成本支出的前提下，提高產品測試的正確率，贏得客戶的信任以及源源不斷的訂單。

資料探勘在經過多年的演進，配合電腦硬體的突飛猛進、資訊技術及新演算法的開發，已從學界及實驗室的純理論，漸漸轉變成為具有實用性的科技。當競爭對手在晶圓測試技術及統計資料分析方面都具有相當能力時，應用資料探勘技術所獨有「從組合在一起的資料中，發現出隱藏價值的知識，以提供給企業專業人員參考」[8]的優點，將能取得其他人所沒有的領先優勢。

本論文提出一個運用資料探勘技術於晶圓測試資料分析的架構；透過與業界專家一同開發的偵測演算法，結合實際的晶圓針測資料為實驗數據，對晶圓測試的誤宰特徵進行分析，再應用資料探勘的分類技術（Naïve Bayes、C4.5決策樹、倒傳遞神經網路以及支持向量機）檢測其準確度；從實驗所得到超過 80% 的準確率，可以看出這些偵測演算法，確實對於針測誤宰的特徵發現，具有相當良好的效果；同時也從實驗中的 C4.5 決策樹推論法，發掘出可針對誤宰現象進行快速判斷的決策樹規則。

實驗二以資料探勘技術中的關聯規則法，對晶圓測試時的環境及設定參數進行分析，嚐試找出與誤宰狀況有相關的原因所在，並將所推論出來的結果，提供給業界作為參考資訊後，透過業界後續的驗證程序，證明關聯規則所推論出來的結果，具有相當的正確性。

未來研究方向：

1. 在實驗一所得的各項分類技術數據中，可以看出 C4.5 決策樹具有相對高的準確度，同時分析數據時的效能也較佳；但因為取樣的產品為邏

輯 IC 類，若能再將其他類產品（如類比 IC 或記憶體 IC）也加入分析，或許會發現其他的分類技術會更適用。

2. 若能將決策樹所推論出的規則內建在業者的生產流程中，當機台端的針測資料產生並上傳到系統之後，便可由系統依據決策樹規則進行計算，而達成即時偵測及決策的功能，減少誤宰現象的連續發生。由於業者所承接客戶的產品種類高達數百種，C4.5 決策樹可先針對每個產品的歷史資料計算出決策樹規則，並預儲在資料庫中，之後便不須在每次要判斷誤宰狀況時，都必須再去存取歷史資料庫，這樣的做法，不僅可以讓業界較易了解其決策規則，而且可以有較佳的效能。
3. 實驗一中，看到倒傳遞神經網路雖然分析數據所耗費的時間較多，但卻具有最高的準確度；若與決策樹規則配合，則可當成第二道的分析工具，當決策樹規則發現可能有誤宰現象時，再轉給倒傳遞神經網路進行更進階的計算，以求能有更準確的決策依據。
4. 本論文中所建立的實驗架構，也可以運用在其他不同領域上，例如：透過分析員工的出勤記錄、活動參與記錄、考核記錄及薪資結構等等資訊，搭配離職員工的歷史記錄，也可建立出一些「偵測演算法」，並透過分類技術進行分析，最後建立出可偵測員工離職可能性的決策樹規則。

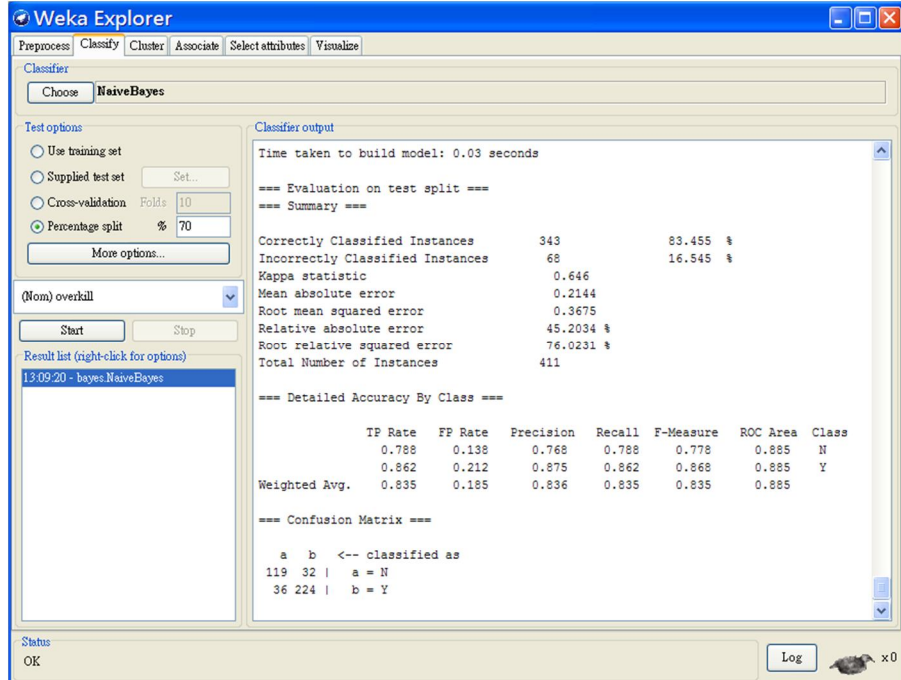
目前半導體業界對於資料探勘的技術導入，通常都用於資料倉儲建置以及數據的分析，對於生產線上的即時偵測應用則較少著墨，我們期待未來的研究可將這些方法，以更實用的方式應用在業界中，發揮其效用，甚至再從這些實務的應用中，研發出更新、更有效的資料探勘技術。

參考文獻

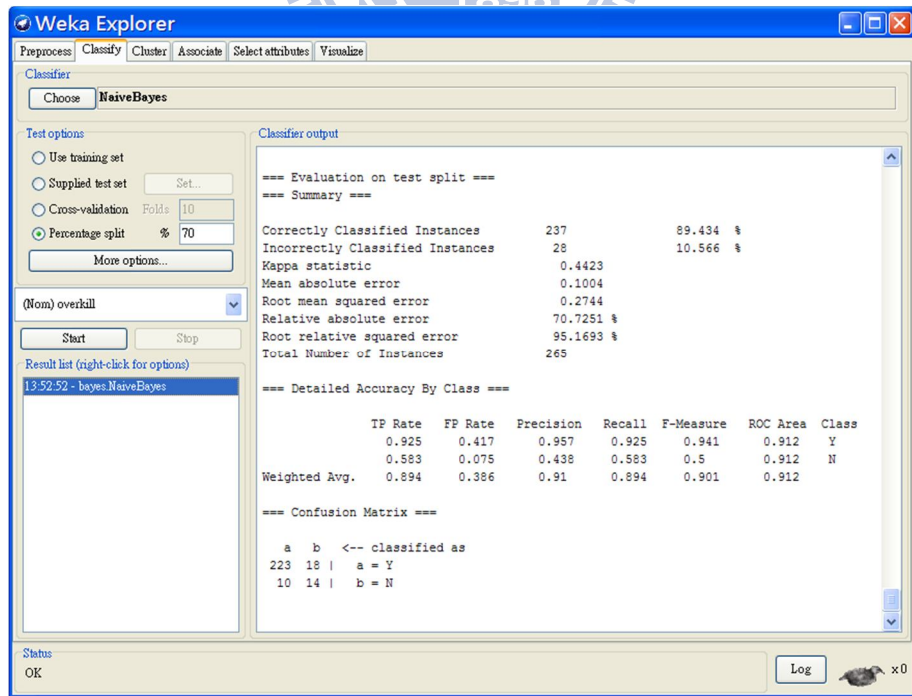
- [1] Agrawal et al., “Mining association rules between sets of items in large database”, Proceedings of the ACM SIGMOD International Conference on Management of Data, Washington, D.C., United States, pp.207-216, 1993
- [2] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules in large database”, Proceedings of the 20th International Conference on Very Large Data Bases, Santiago de Chile, Chile, 1994
- [3] T. Bayes, “An essay toward solving a problem I the doctrine of chances”, Philosophical Transactions of the Royal Society of London, vol.53, pp.270-418, 1764
- [4] Michael J.A. Berry; Gordon S. Linoff, 吳旭志，賴淑貞譯，”資料探礦理論與實務－顧客關係管理的技巧與科學”，維科圖書有限公司，2001
- [5] C. Cortes and V. Vapnik, "Support-Vector Networks", Machine Learning, vol.20, pp.273-297, 1995
- [6] W.J. Frawley, G. Piatetsky-Shapiro and C.J. Matheus, “Knowledge discovery databases: An overview, in Knowledge Discovery in Databases” (G. Piatetsky-Shapiro and W. J. Frawley, eds.), Cambridge, MA: AAAI/MIT, 1991-1-27.
- [7] J. Joyce, “Baye’s Theorem”, The Stanford Encyclopedia of Philosophy (Fall 2003 Edition), Edward N. Zalta (ed.), 2003
- [8] C. Kleissner, “Data mining for the enterprise” in Proc. of the Thirty-First Hawaii Int. Conf. on System Sciences, Kohala Coast, USA, vol.7, pp.295-304, Jan. 1998
- [9] D.B. Parker, ”Learning Logic Technical Report TR-47, Center for Computational Research in Economics and Management Science”, MIT,1985
- [10] J.R. Quinlan, “Induction of Decision Tree”, Machine Learning, Vol.1, pp.81-106, 1986
- [11] J.R. Quinlan,, “C4.5: Programs for Machine Learning”, Morgan Kaufmann, San Diego, 1993
- [12] J.R. Quinlan, “Improved use of continuous attributes in C4.5”, Journal of Artificial Intelligence Approach, vol.4, pp.77-90, 1996
- [13] D. Rumelhart, G. Hinton and R. Williams, ”Learning Internal Representations by

- Error Propagation”, Parallel Distributed Processing : Foundations , pp.318-362, MIT Press, 1986
- [14] V. Vapnik, “The Nature of Statistical Learning Theory”, Springer-V verlag, NY, USA, 1995
- [15] P. Werbos, "Beyond Regression: New tools for Prediction and Analysis in the Behavioral Sciences", PhD thesis Harvard,1974
- [16] WEKA software. URL: <http://www.cs.waikato.ac.nz/ml/weka>, 2007
- [17] 李俊宏, 古清仁 「類神經網路與資料探勘技術在醫療診斷之應用研究」, 工程科技與教育學刊第七卷第一期, 2010
- [18] 吳永安, 「即時預測模型於晶圓測試回收良率之應用」, 國立交通大學理學院應用科技碩士論文, 2011
- [19] 吳瑞崎, 「運用資料探勘技術建構半導體封裝業之品質改善系統」, 國立交通大學資訊管理研究所碩士論文, 2006
- [20] 周嵩能, 「以微網誌語料進行情緒辨識之研究」, 國立台南大學資訊工程學系碩士論文, 2010
- [21] 祝健華, 「建構電腦整合製造決策支援離型系統一個案研究」, 國立交通大學管理學院資訊管理碩士論文, 2004
- [22] 陳品潔, 「應用資料探勘技術於半導體業製程良率改善之研究」, 國立交通大學管理學院碩士論文, 2004
- [23] 黃振聲, 「應用資料探勘技術於智慧型電腦整合製造系統建構之研究—以 IC 封裝測試業為例」, 國立交通大學資訊管理研究所碩士論文, 2005

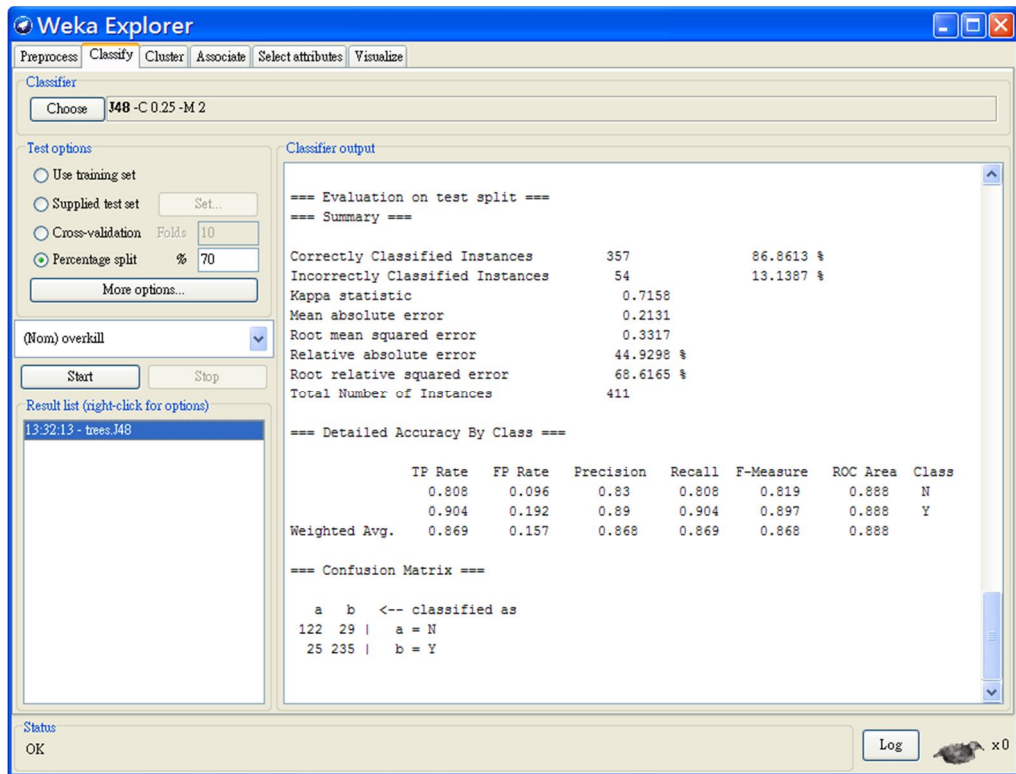
附 錄



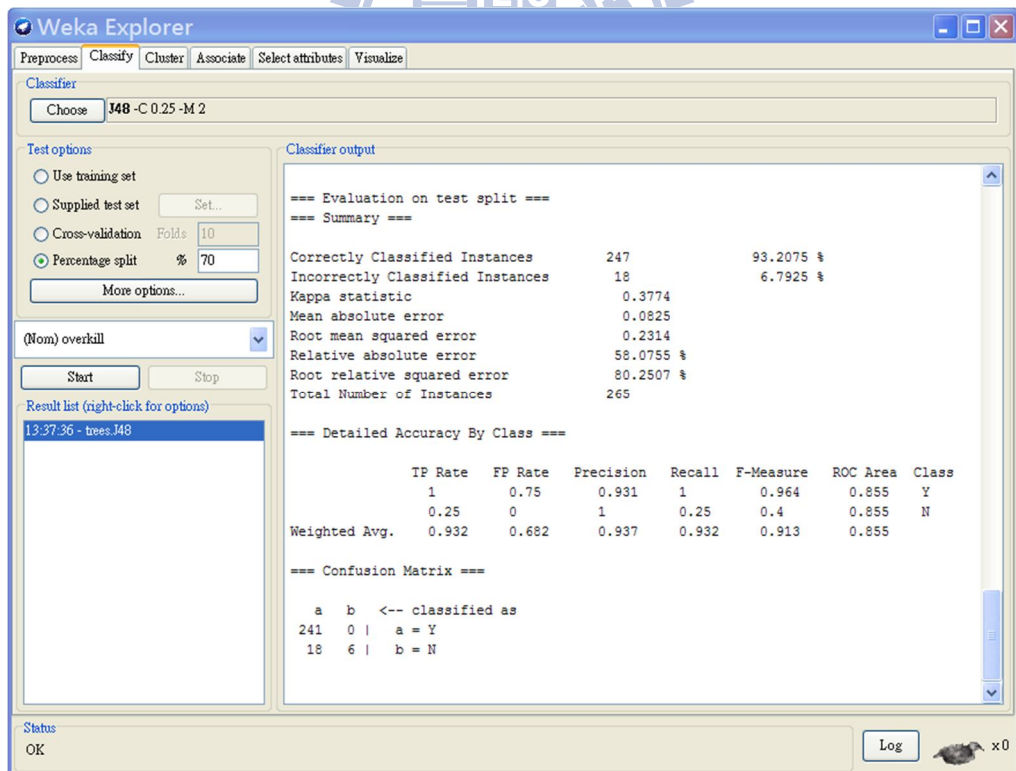
附錄圖 1：A 產品 Naive Bayes 演算法執行結果



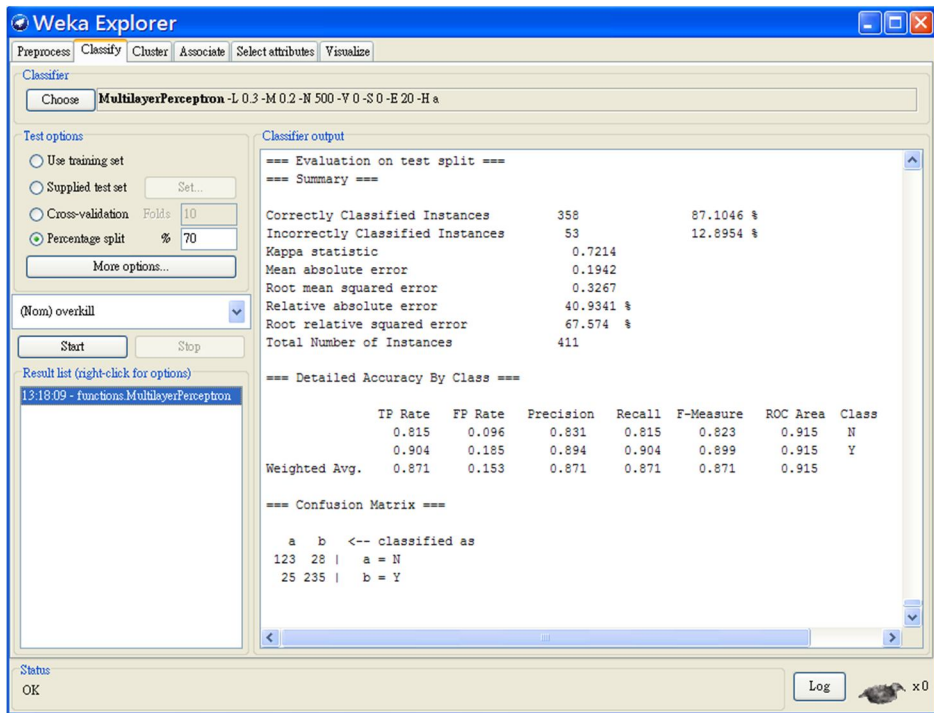
附錄圖 2：B 產品 Naive Bayes 演算法執行結果



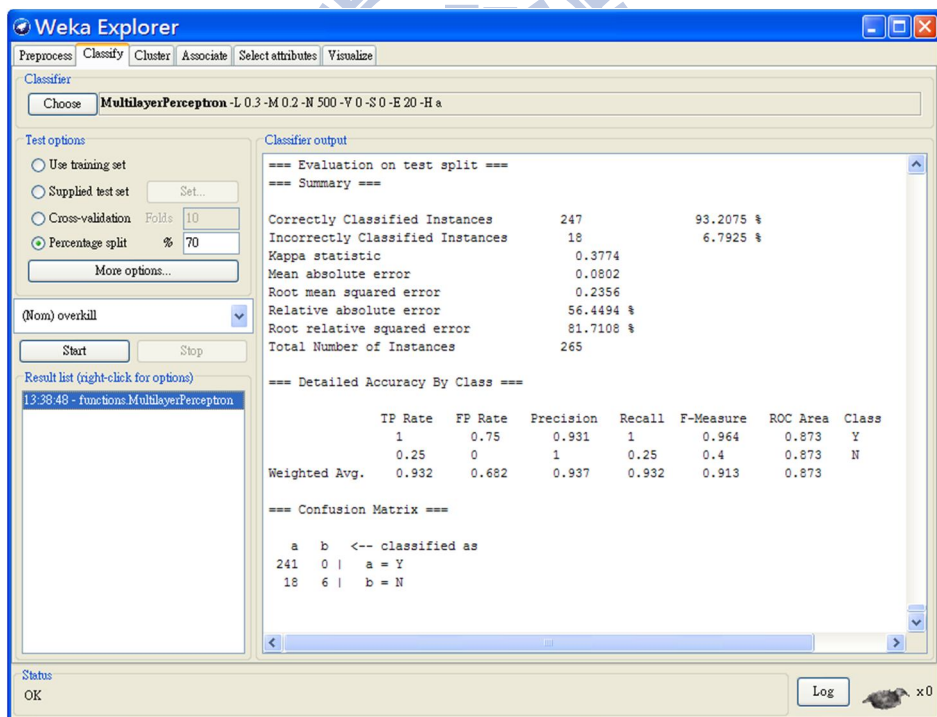
附錄圖 3：A 產品 Decision Tree(J48)演算法執行結果



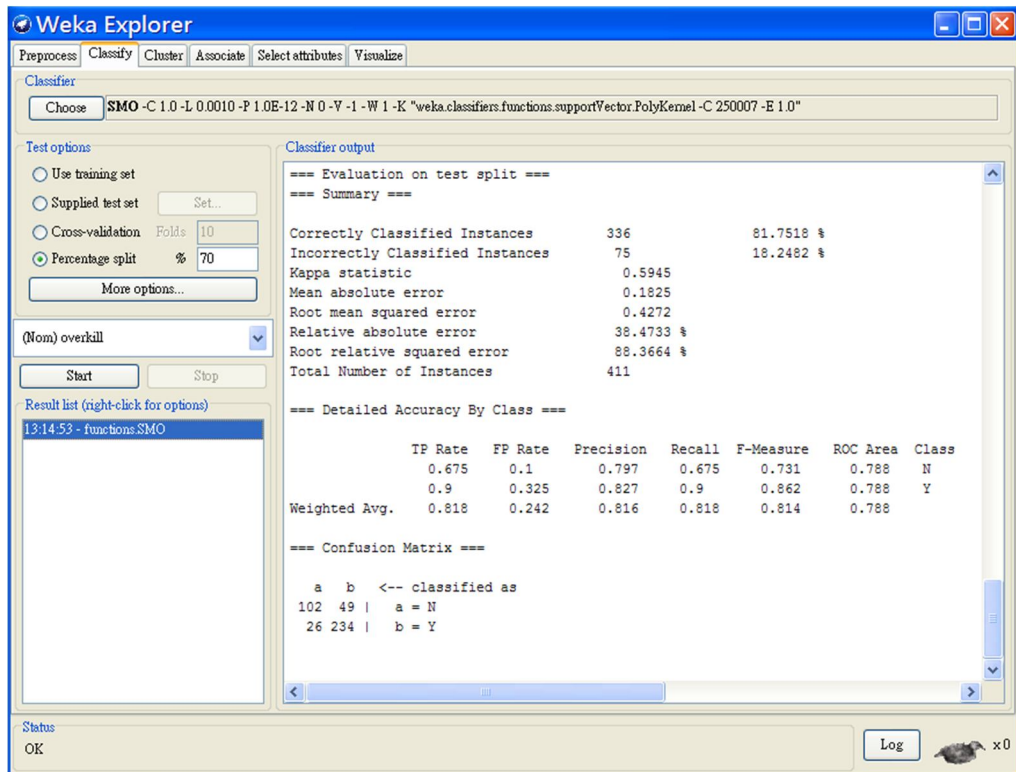
附錄圖 4：B 產品 Decision Tree(J48)演算法執行結果



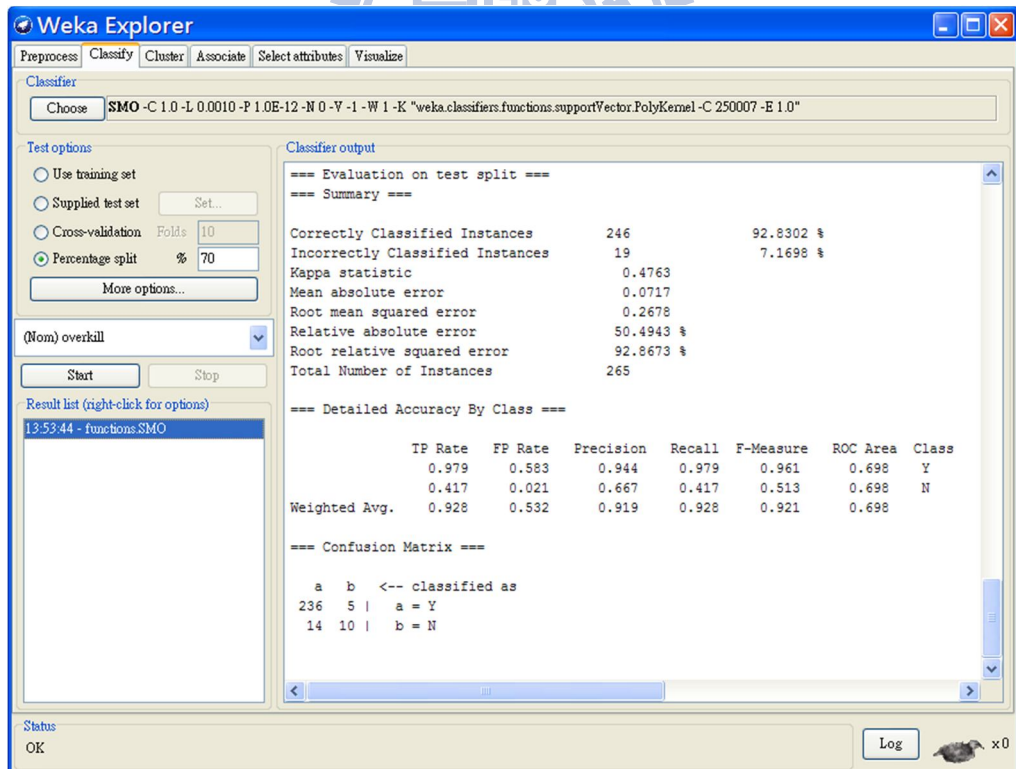
附錄圖 5：A 產品 MultilayerPerceptron 演算法執行結果



附錄圖 6：B 產品 MultilayerPerceptron 演算法執行結果



附錄圖 7：A 產品 SVM(SMO)演算法執行結果



附錄圖 8：B 產品 SVM(SMO)演算法執行結果