

運用資料探勘技術於藍圖材料表校對自動化之研究

孫懷谷¹ 王震宇² 鄧世剛¹ 雷賀君^{3*}

¹ 國防大學理工學院動力暨系統工程學系

² 中山科學研究院航空研究所

³ 國防大學理工學院國防科學研究所

摘 要

航太產業相較於其他產業，對於藍圖材料表之校對更為謹慎與重視，為避免因藍圖資料闕漏而造成不必要之成本負擔，本研究選用數種資料探勘技術於藍圖材料表校對之適用性進行探討。其主要目的為透過資料探勘技術從藍圖材料表中找到藍圖校對的規則，進而取代現行人工藍圖校對作業，並探討適用於藍圖材料表校對自動化之資料探勘技術。本研究透過資料探勘方法找出各欄位之間之關聯性，發展出藍圖校對規則，並從中評選出較適合之資料探勘方式，以提供校對人員作為後續運用於藍圖材料表自動化校對時之參考依據。結果顯示運用關聯規則之 Apriori 演算法可以得到比其他資料探勘模型更好的規則庫，以提供後續自動化校對之標準。

關鍵字：資料探勘、藍圖校對自動化

The Study of Data Mining to Automation Checking Rule of Blueprint

Huai-ku Sun¹, Jerry Wang², Shigan Deng¹, and Ho Chun Lei^{3*}

¹Department of Power Vehicle and Systems Engineering, Chung Cheng Institute of Technology, National Defense University

²Department of Aeronautical, Chung-Shan Institute of Science and Technology, Armaments Bureau, M.N.D

³Graduate School of Defense Science, Chung Cheng Institute of Technology, National Defense University

ABSTRACT

Aerospace industry has more cautious and attention on checking material form of blueprint than other industries. To avoid unnecessary cost caused a blueprint data leakage. This research used various data mining methods to analyze the checking rule for material list of blueprint. The purpose of study is to gain blueprint checking rules from data of material list to replace manually checking with automatically checking. The relationships between columns of database and expanding the checking rules of blueprint were found out. And the suitable data mining method was proposed to check material list of blueprint for checker in future. The results showed that the checking accuracy of Apriori algorithm is better than other model, reflecting a satisfactory accuracy of the blueprint checking by association rule.

Keywords: data mining, blueprint automation checking rule

一、前言

在產品設計開發過程中，藍圖校對工作是一項不可忽視的工程，因設計藍圖的失誤，會造成後續製造階段與整個生命週期的時程延誤與成本損失。航太產業對於藍圖材料表的校對尤其重視，所有校對工作都要遵照標準作業程序，更有專門的校對員進行藍圖校對的工作。航太產業的藍圖校對員多是由經驗豐富的繪圖員晉升以確保藍圖校對工作的品質。由於飛機零件造價不菲且具有其獨特性，藍圖校對的疏失，會導致圖面資料與材料清單資料不一致，而造成設計與製造時程與成本的損失。若是在設計階段能進行精準的藍圖校對，更能節省不少成本。

現階段藍圖校正工作執行方式仍根據校對人員的經驗判斷或相關標準手冊以人工徒手方式進行各項校正，所耗費的人力工時相對較高，且容易因一時疏忽導致校對錯誤的情況產生，進而降低產品可靠度。由於過去工程圖多採用人工繪圖，當電腦輔助設計軟體開始取代人工作業後，學者們紛紛針對如何將過去人工繪製之工程圖轉為電子檔，以及針對圖形識別來判斷原始手繪圖檔與轉換後之差異進行研究。陶以誠、蔡啟良 [1][2] 就曾建立一個以知識為基礎的工作藍圖讀圖系統，將工程圖的線條資料分成幾何與非幾何兩種。主要工作是掃描圖形後再透過影像處理的方式擷取工件圖的資料並建立介面表，最後利用介面表進行資料推論以檢查工件圖的正確性與完整性。其他針對工件圖所處理的問題大多都是利用掃描器或相機等影像擷取設備將工件圖數位化，再透過影像識別方式進行圖形重繪，並與原圖進行比對[3][4]。

製造業應用資料探勘技術始於 1990 年代，之後便逐漸為之接受。其最早應用於預測產品之維持及失效週期，之後逐漸運用於設計、生產、品管排程甚至是決策系統之支援等。透過資料探勘，製造業可迅速分辨製造過程中隱含之模式，進而決定或改善其產品品質，以期增加利潤及競爭性。

大部分之資料探勘研究均專注於增進演算法與應用程式之工作效率，而非提高其介面之適用程度。Neaga 與 Harding [5] 提出一整體性的方法可將各種資料探勘技術應用於製造業，並建立適合工業運用之方法理論及軟體工

具。其主要為建構一整合企業體中多種應用程式間複雜介面之軟體架構(Framework)，此方法清楚定義一通用之企業知識模組，其資料探勘模組可立即沿用企業以往之資料庫，使得其適用性更勝一籌。

在機械工程設計上，由於其具有多重規則、多維度及決策流程非線性等特性。因此，資料探勘在此一領域之研究相對於其他領域而言，發表文件較少。但近年因線上即時同步作業之發展趨勢，設計與製造兩大部門以往之作業流程順序亦因此漸趨同步化。機械工程設計也因之成為資料探勘之重要探討領域。生產流程中，以往之設計習慣為在設計階段思考產品之生產流程，並依據特定製造流程或技術，設計並輸入特定之限制條件。因此，其有非常大的發展空間可供資料探勘改良其原有之產品設計、生產流程，並萃取其中未曾發現之隱含知識。

Sim 與 Chan[6]發展出一套專家知識系統，可自動篩選產品。所使用的為依據生產者提供之生產經驗所訂定符合產品最佳化規則之階層式知識結構。Ishino 與 Jin[7]運用資料探勘技術，自設計階段使用之 CAD 系統產品檔案中，獲得設計所需之知識模式。Romanowski 與 Nagi[8]運用資料探勘分析一般材料清單 (GBOMS, Generic Bills of Materials)，並自同一系列之類似產品設計中找出其類似之設計處，亦可找出相異的設計配置。

然而，電腦輔助設計軟體的發展，幫助繪圖員能在電腦上繪製各式 2D 或 3D 圖檔。繪製好的工程圖卻仍是輸出成紙本後交給校對人員進行藍圖校對。藍圖校對的內容可分為圖型校對與文字校對。其中圖型校對包含：不同零件圖之尺寸是否合乎公差要求、組裝時是否會出現干涉等問題。而文字的校對包含有：圖面上的註釋以及藍圖材料表的校對。

本研究主要之研究方向就是針對工程藍圖上的材料表進行校對。一般材料表所包含的資料內容有：圖號、零件件號或編號、零件名稱及規格、圖域、材料、規範等。主要校對的內容包含材料與材料相對應的材料規範是否正確、特殊加工程序是否正確。搜尋國內外相關文獻發現，目前在藍圖的審查上，尚未有自動化的輔助工具。因此本研究選用數種資料探勘技術於藍圖材料表校對之適用性進行探

討。其主要目的在透過資料探勘的技術，從大量且雜亂的藍圖材料表中建立其規則性，進而取代現行藍圖校對方式。並從中評選出較適合之資料探勘方式，以提供校對人員作為後續運用於藍圖材料表校對自動化時之參考依據。

二、文獻探討

資料探勘為一以計算機智慧(CI, Computational Intelligence)為主之新技術與理論，運用目前電腦的快速處理速度及各種不同的啟發式演算法，對龐大數量中不同類別、模糊、缺項的資料進行資料探勘。其主要目的為透過分析龐大的資料庫，挖掘出具有高價值之隱含意義，以支援企業決策。適用之資料型態可以是文字、數據、圖像、甚至是網路多媒體均可進行資料探勘。目前發展較健全之資料探勘方式為針對資料庫、倉儲等表格化資料進行探勘之相關技術。

2.1 關聯規則-Apriori演算法

關聯法則因不需先行建立分類規則，故有利於發覺隱藏之關聯資料。目前除最早發展之商業用途外，醫學上亦多所運用，如以病人病徵之關聯判定可能之疾病並分析最佳之療程或療法等。而運用關聯法則最有名之案例為超市購物推車分析。其運用消費者購買某商品後，亦會購買另一商品之關聯性，分析出「尿布與啤酒」之關聯性。其中就資料庫而言，於購買紀錄中發覺特定品項間相互關聯性其實並不容易發覺。但運用資料探勘後，即可分析出所有購物紀錄中，被購買商品之特殊關聯性。此後，便可以此為進、存貨與店內商品擺設位置之參考依據。

關聯規則最早是由 Agrawal [9]提出，他對關聯規則定義如下：

假設 $I=\{I_1, I_2, \dots, I_m\}$ ， I 是 m 個商品項目的集合。 $D=\{t_1, t_2, \dots, t_n\}$ ， D 為 n 個客戶交易的總集合。其中， $t_i=\{I_{i1}, I_{i2}, \dots, I_{ik}\}$ ， t_i 代表第 i 位元客戶的交易資料。

關聯規則的代表式為：“If condition then result”，用數學法表示就是 $X \Rightarrow Y$ 。 X 、 Y 就是關聯規則中的項目組。

在關聯規則中有兩項評鑑規則的重要參數：支持度(Support)與信賴度(C Confidence)。

支援度是指 X 項目集合與 Y 項目集合，同時出現在 D 交易總集合的次數，除以 D 交易總集合個數，也就是同時發生 X 、 Y 事件的機率。

信賴度是指 X 項目集合與 Y 項目集合，同時出現在 D 交易總集合的次數，除以 X 項目集合在 D 交易總集合出現的次數，也就是在 X 事件發生的情況下， Y 事件發生的機率。

除此之外，另有學者提出興趣度(Interesting)或稱增益(Improvement)的規則指標。其興趣度計算公式如下：

$$\frac{\text{Confident}(X \Rightarrow Y)}{P(Y)} = \frac{P(X)}{P(Y)} \quad (1)$$

當興趣度大於 1 時，此規則實際意義越高，當規則小於 1 時，則此規則並無太大實際效益。

Apriori 演算法是在關聯法則中最具代表性的演算法，其執行步驟如下。

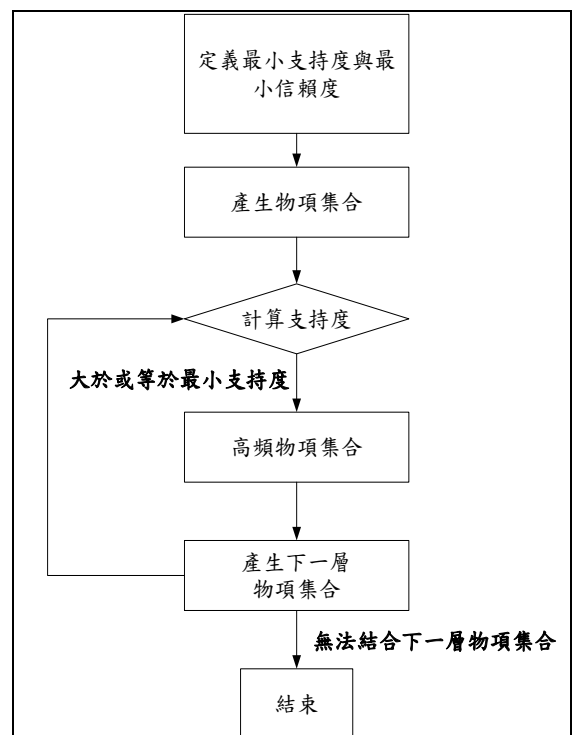


圖 1. Apriori 演算法流程圖

1. 訂定最小支持度與最小信賴度

2. Apriori 法則採用候選物項集合的觀念，先產生物項集合，稱之為候選物項集合，若候選物項集合的支援度大於或等於最小支

持度，則該候選物項集合為高頻物項集合。

3.透過資料庫取出所有交易，得到候選單物項集合(Candidate 1-itemset)的支持度，再找出高頻單物項集合(Large 1-itemset)，利用這些高頻單物項集合的結合，產生候選 2 物項集合(Candidate 2-itemset)。

4.掃描資料庫，得出候選 2 物項集合的支援度，再找出高頻 2 物項集合，並利用高頻 2 物項集合的結合產生候選 3 物項集合。

5.重複掃描資料庫直到不再結合產出新的候選物項集合為止。

2.2 決策樹-C5.0

決策樹由 Friedman [10]等人提出，1980 年代以來就開始發展，是基於樹結構產生分類和迴歸模型的過程，是一種產生二元樹的技術。CART 模型適用於目標變數為連續型和類別型的變數，如果目標變數是類別型變數，則可以使用分類樹(classification trees)，目標變數是連續型的，則可以採用迴歸樹(regression trees)。Han 和 Kamber [11]指出現今決策樹的演算法大致有 C5.0、CART、CHAID 與 QUEST 這四種演算法。CART 與 C4.5/C5.0 演算法的最大相異之處是其在每一個節點上都是採用二分法，也就是一次只能夠有兩個子節點，C4.5/5.0 則在每一個節點上可以產生不同數量的分枝。Perreault 和 Barksdale[12]提出 CHAID(Chi-square automatic interaction detector)分析方法，CHAID 是由 AID 演變而來的。CHAID 會防止資料被過度套用並讓決策樹停止繼續分割，依據的衡量標準是計算節點中類別的 P 值大小，以此決定決策樹是否繼續分割，所以無需要做樹剪枝。Quinlan[13]提出 C4.5/C5.0 演算法，C5.0 演算法是由 C4.5 演算法所演變而來。C5.0 不限制只能做二元分割，這是與其它決策樹分類法不同的地方 [14]。

優點：

C5.0 模型在面對資料遺漏和輸入欄位很多的問題時非常穩健。

C5.0 模型通常不需要很長的訓練次數進行估計。

C5.0 模型比一些其他類型的模型易於理解，模型推出的規則有非常直觀的解釋。

C5.0 也提供強大的增強技術以提高分類

的精度。

2.3 神經網路

神經網路的發展約從 1940 年開始至今，已經在許多不同的領域被廣泛的應用。所謂的神經網路就是利用電腦來模仿生物神經網路的處理系統。神經網路主要可分為監督式學習與非監督式學習兩種。監督式學習就是以輸入與輸出的資料去找尋內在對應的規則或關聯；非監督式學習則是在只有輸入資料時去學習內在的聚類規則。監督式的神經網路包含倒傳遞網路(Back Propagation Network)、多層函數連結網路(Multi-layer Functional Link Network)、通用迴歸網路(General Regression Network)、學習向量量化網路(Learning Vector Quantization Network)、半徑式函數網路(Radial Basis Function Network)等。其中最常被使用的監督式網路為倒傳遞網路。P. Werbos 於 1974 年在其博士論文中提出了隱藏層的學習演算法，這是最早已知的倒傳遞神經網路模型 [15]。倒傳遞神經網路的架構如圖 2 所示，包含輸入層、隱藏層及輸出層，而隱藏層的數目可以不只一層。倒傳遞網路是利用輸入向量和與輸入向量相對應的目標向量來訓練網路，直到網路能經由適當的方式來分類輸入向量。標準的倒傳遞演算法是一種梯度下降(Gradient Decent)演算法，網路權重值是沿著性能函數為負的梯度方向移動。網路中的神經元最常用的非線性轉換函數為雙彎曲函數

(sigmoid function, $f(x)$)，其中 $f(x) = \frac{1}{1 + e^{-x}}$ ， $f(x)$ 的值介於(0,1)之間，當 X 趨近於正負無窮大時， $f(x)$ 趨近於 0 或 1。

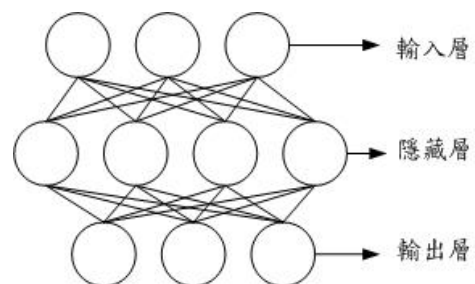


圖 2. 倒傳遞神經網路架構

三、研究方法

由於過去學者研究中，從未針對工作藍圖材料表的校對工作進行研究。本研究期望透過資料探勘技術進行工作藍圖材料表的自動化校對，因此選擇了 Apriori 演算法、C5.0 及倒傳遞神經網路三種方法進行實證與探討。

以本研究之主要方向而言，材料屬性表可透過整理後，使用一般資料探勘方法建立材料屬性表的規則。本研究目的在找出藍圖材料表自動化校對所適用之資料探勘技術，將藍圖材料表資料輸入資料庫中，透過數種資料探勘找出材料表欄位之間的相關規則並針對不同模型進行比較。由於在藍圖材料表校對上，材料與加工法則多具有高度關聯性或是具有因果關係。因此本研究選用關聯規則、決策樹與神經網路作為研究之方法，以產出藍圖材料表之規則供自動化校對程式參考。研究之架構圖如圖 3 所示。

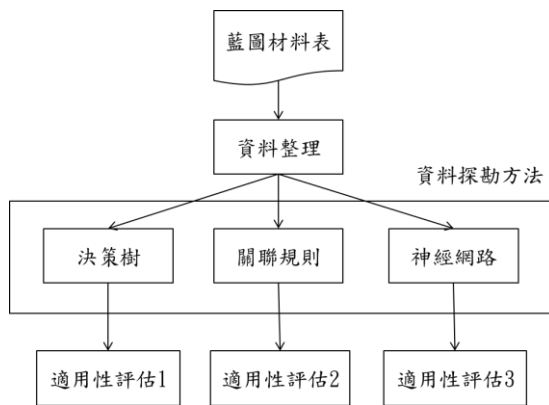


圖 3. 研究架構

在過去研究中，少有人將三種演算法合併比較。但在本研究中考量藍圖材料表的資料形態，可透過關聯規則尋找出校對規則，也可能從決策樹找出規則或是將欄位進行適當分類，神經網路同樣用來找尋欄位之間關係的強弱來進行分類。分類出的結果同樣可以轉成校對的規則。因此針對三種不同的演算法進行實證，期望找出較適合藍圖校對的探勘技術。

3.1 欄位資料整理

資料探勘所採用的技術方法往往會因資料型態、內容的不同而有所改變。因此不同類型的資料所適用之方法也不盡相同。工作藍圖材料表所記錄的資料欄位包含有圖號、零件件號或型號、零件名稱及規格、材料、規範、圖

域、材料尺寸 1、材料尺寸 2、材料尺寸 3、原料等。

其中圖號為每個工作藍圖之識別號，圖域是指該零件或部件圖位於圖紙上的位置，零件件號與型號在藍圖材料表中所記錄的是該張藍圖所使用的第幾個零件，與圖號相同，僅做為每個零件之識別號，因此圖號與圖域及零件件號三個欄位在資料探勘過程中並不具有實質意義，故先行剔除。材料尺寸 1、材料尺寸 2、材料尺寸 3 三個欄位在藍圖材料表中代表的是原料的尺寸大小，但在過去藍圖材料表中並無詳實記載每一個工作料件之原型尺寸大小，缺漏資料甚多，無法透過資料探勘技術獲得材料尺寸與其他欄位間是否存在關連性。

所以本研究僅針對零件名稱及規格、材料、與材料相對應之規範及原料四個欄位進行資料探勘，期望找出四個欄位間是否存在一定的規則或關連性能夠作為藍圖材料表校對之依據。

3.2 Apriori 參數設定

Apriori 主要用來尋找資料間是否存在某種關連性與規則性，若藍圖材料表間具有零件名稱=A1，則材料=B1 的規則形態，就可以用來檢驗藍圖材料表上的資料是否有誤。

因此針對 Apriori 演算法，我們設定每一個欄位皆可以為輸入欄位也可以為輸出欄位，因為單方向的搜尋，可能會遺漏某些部份的規則。但在藍圖材料表校對中，講求的是高正確率，因此只要是曾經產生過的正確規則，都希望能被保存做為日後校對之依據。因此對於 Apriori 所產生出的規則我們設定最小支持度為 0%，最小信心度為 80%。

3.3 決策樹參數設定

決策樹的規則產生是以其中一個欄位為輸出欄位，其他欄位為輸入欄位來產生規則。其建立之規則可能具有較長的長度，如 A=A1 and B=B1 and C=C1, then D=D1。因此本研究設定產生規則的長度為 2，及所產生出之規則形態為 If A=A1, then D=D1。將藍圖材料表所要探勘的四個欄位進行四次規則搜尋，每次設定其中一個欄位為輸出欄位，其餘三個欄位為輸入欄位。

3.4 神經網路參數設定

將資料隨機以 60%、20%、20% 的比例分配為訓練資料、測試資料、驗證資料。訓練神經網路架構之停止機制分為兩種，第一個是將停止機制設為 3000 個循環即停止，輸入層神經元個數 17 個，輸出層神經元個數 20 個，由神經網路訓練找出隱藏層層數與各層最佳之神經元個數。第二個停止機制是設定在收斂至精確度為 80%，同樣設定輸入層神經元個數 17 個，輸出層神經元個數 20 個。由神經網路訓練找出最佳隱藏層層數與各層神經元個數。

四、實驗結果與分析

本研究擷取 38 個工作藍圖，共 103 筆零件資料來進行案例模擬。每一筆零件資料所採用的資料欄位為：零件名稱及規格、材料、規範原料。整理過後的藍圖材料表欄位及內容如表 1 所示。

其中零件名稱及規格在 103 筆資料中共有 57 種不同零件，使用的材料有 15 種，所採用的規範有 20 種，其原料有 21 種。

表 1. 工作藍圖材料表

零件名稱及規格	材料	規範	原料
SLINGER	17-4PH	AMS5643	硬圓管
FRONT FRAME ASSY	NA	NA	組件
FRONT FRAME (INVEST, CASTING)	17-4PH	AMS 5355	鑄料
INSERT	NA	NA	嵌入件
PIN	NA	NA	標準件
INSERT	NA	NA	嵌入件
INSERT	NA	NA	嵌入件
FITTING	AN816-6J	REWORK	零件
JET	NA	NA	板
BODY	17-4PH	AMS5643	板
TUBE	17-4PH	AMS5643	軟圓管
PLUG	17-4PH	AMS5643	塞
SLEEVE	CRES 304	AMS5639	套管
SHROUD	NA	NA	罩蓋
BODY	CRES 304	AMS5639B AR	桿

4.1 Apriori 驗證

本研究首先將工作藍圖材料表之四個欄位同時設定為輸入及輸出欄位，最小支持度設定為 0%，最小信心度設為 80%，如表 2 所示。

表 2. Apriori 參數設定

筆數	103 條
輸入欄位	4
輸出欄位	4
最小支持度	0%
最小信心度設	80%

在此規則條件內，運用 Apriori 方法計算結果如表 3 所示。Apriori 此方法一共產生 833 條規則，其中規則的最小支持度為 0.971%，規則的最大支持度為 25.243%。結果顯示本研究所採用的藍圖材料表所包含之資料多樣而重複性低。而 Apriori 所找出之規則最小信心度介於 80%~100% 之範圍內，其中有 816 條規則之信心度均為 100%。且其中有 816 條規則之信心度均為 100%。最後本研究將所產生之規則庫與原始 103 筆資料進行校對驗證，其對於原始資料判斷之準確度高達 97.95%。

表 3. 計算結果

規則總數	833 條
最小支持度	0.971%
最大支持度	25.243%
最小信心度	80%
最大信心度	100%

4.2 決策樹驗證

本研究運用決策樹進行資料探勘時，設定產生規則的長度為 2。規則搜尋時，本研究將工作藍圖材料表中四個欄位，選擇三個欄位設定為輸入欄位，另一個欄位則設為輸出欄位，重複此方式四次進行資料探勘，如表 4 所示。例如在相同的資料集合將規範當作規則輸出欄位，零件名稱及規格、材料與原料當作輸入欄位。本研究將資料集分成四個子集合進行交叉驗證後，可得到十組規則集合。每一組規則集合中包含 15 至 26 條規則不等。最後將產生之規則庫與原始 103 筆藍圖材料表資料進行驗證，其對於原始資料判斷之準確度為 88%。

表 4. 決策樹參數設定

筆數	103 條
規則長度	2
輸入欄位	3
輸出欄位	1

4.3 神經網路驗證

我們將同樣的案例資料以神經網路進行資料探勘尋找規則。由於神經網路一直被視為是黑盒子，無法明確解釋其代表的規則與意義。僅能將測試的規則丟入神經網路，由網路架構來判斷是否正確。本研究在設定停止機制為 3000 個循環下，訓練過後的最佳神經網路架構之輸入層具有 17 個神經元，隱藏層為 3 個神經元，輸出層為 20 個神經元，如表 5 所示。但將驗證資料投入經過訓練後得到的神經網路。可得到神經網路對於資料的推估準確度僅有 32.394%。此外，本研究重複進行運算神經網路進行資料探勘尋找規則，將停止機制收斂在精確度 80% 的最佳神經網路架構包含兩層隱藏層，其輸入層具有 17 個神經元，第一層隱藏層為 11 個神經元，第二層隱藏層為 10 個神經元，輸出層為 20 個神經元，如表 6 所示。而將驗證資料投入此神經網路架構後，得到此網路架構對資料推估的準確度為 51.456%，本研究嘗試將訓練神經網路的準確度提高，但造成神經網路無法收斂的情況。

表 5. 神經網路參數設定

筆數	103 條
停止機制	3000 次
輸入層	17
隱藏層	3
輸出層	20

表 6. 神經網路參數設定

筆數	103 條
停止機制	精確度 80%
輸入層	17
第一隱藏層	11
第二隱藏層	10
輸出層	20

4.4 小結

綜整 Apriori、C5.0、神經網路三項研究方法之準確率如表 7 所示。結果顯示在藍圖材料表的校對上，Apriori 方法準確率較 C5.0 與神經網路精準。

表 7. 神經網路參數設定

方法	準確率
Apriori	97.95%
C5.0	88%
神經網路	51.456%

五、結論與未來研究方向

由於航太產業的特殊需求，提高藍圖材料表校對之準確率與效率以及減少藍圖校對材料表之人力工時，對於航太產業來說有很大的助益。經研究分析後可得到以下結論：

1. 因為本研究針對航太產業在實際藍圖材料表上具有資料重複性低、樣本少等特徵，因此在神經網路的訓練成效並不顯著。

2. 比較決策樹與 Apriori 之結果，雖然透過決策樹可以產生較精簡之規則。但是運用在藍圖材料表校對上仍顯不足。

3. Apriori 產生之規則雖然高達 833 條，但對於要求正確率要高的藍圖材料表校對來說，其準確率優於其他二者，也能正確校對更多樣的材料表欄位。由於產業特性，隨著時間增加，可以提高規則的支持度與信心度，且規則數並不會大幅增加。

4. 由於目前並無相關藍圖材料表之資料庫，因此在第一次進行校對時，須先用人工將現有材料表之規則建立在資料庫當中。

5. 根據未來藍圖材料表的資料增加及欄位的變化，將增加特徵自動選取機制，來判斷適合探勘之資料欄位。

6. 未來將利用 Apriori 演算法，建立自動藍圖材料表校對機制，將現行人工校對改為自動化校對機制。

根據本研究之結果建議，欲將資料探勘方法運用於藍圖材料表之自動化校對上，建議使用 Apriori 演算法可以得到較好的規則。更可以達到降低藍圖材料表校對工作所耗費之人力工時的目標。

致 謝

本研究由 中山科學研究院 航空研究所
提供相關數據與經費支持，謹此致謝。

參考文獻

- [1] 陶以誠，一個以知識為基礎的工件藍圖讀圖系統。第一部份：自動讀圖與檔案建立，碩士論文，國立交通大學計算機工程研究所，新竹，1987。
- [2] 蔡啟良，一個以知識為基礎的工件藍圖讀圖系統。第二部份：一致性檢驗，完備性檢驗和資料推論，碩士論文，國立交通大學計算機工程研究所，新竹，1987。
- [3] 賈季平，工件藍圖之自動判譯與交互式構圖，碩士論文，國立交通大學計算機工程研究所，新竹，1986。
- [4] 陳啟彰，機械 CAD 資料檔的自動檢驗與錯誤標示，碩士論文，國立交通大學計算機工程研究所，新竹，1987。
- [5] Neaga, E. I. and Harding, J. A. “An enterprise modeling and integration framework based on knowledge discovery and data mining”, *International Journal of Production Research*, 43(6), 1089 – 1108, 2005
- [6] Sim, S. K. and Chan, Y. W., “A Knowledge-Based Expert System for Rolling-Element Bearing Selection in Mechanical Engineering Design,” *AI in Engineering*, Vol. 6, No. 3, pp.125–135, 1992.
- [7] Ishino, Y. and Jin, Y., “Data Mining and Knowledge Acquisition in Engineering Design,” *Data Mining for Design and Manufacturing: Methods and Applications*, D. Braha, ed., Kluwer Academic, Dordrecht, pp. 145–160, 2001.
- [8] Romanowski, C. J. and Nagi, R., “A Data Mining for Knowledge Acquisition in Engineering Design,” *Data Mining for Design and Manufacture: Methods and Applications*, D. Braha, ed., Kluwer Academic, Dordrecht, pp. 161–178, 2001.
- [9] Agrawal, R., Imielinski, T., and Swami, A., “Mining association rules between sets of items in large database,” *Proceeding of the 1993ACM SIGMOD Conference*, Washington, DC, 1993.
- [10] Friedman, H. H. and Friedman, L. “Endorser Effectiveness by Product Type,” *Journal of Advertising Research*, Vol. 19, No. 5, pp. 63-71, 1979
- [11] Han, J. and Kamber, M., *Data Mining: Concepts and Techniques*, Morgan Kaufmann, New York, 2001.
- [12] Perreault, W. D. and Barksdale, H. C., “A Model-free Approach for Analysis of Complex Contingency Data in Survey Research,” *Journal of Marketing Research*, Vol. 7, No. 4, pp. 503-515, 1980.
- [13] Quinlan, J. R. “C4.5 Programs for machine learning,” Morgan Kaufmann Publishers, San Mateo, California, 1993.
- [14] 郭信宏，應用資料探勘技術於面板檢測實證研究，碩士論文，國立中央大學工業管理研究所，桃園，2008。
- [15] 王進德、蕭大全，類神經網路與模糊控制理論入門，第 23-36 頁，全華科技圖書，2002