財金論文叢刊
2014 年 6 月，第二十期，14-32

加值型營業稅選案研究-以財政部中區國稅局某稽徵所爲例

A Study of Value-Added Tax -The Cases of National Taxation Bureau in the Central Area, Ministry of Finance

游崇熙（Chong-Si You）
財政部中區國稅局稅務員

摘要

本研究以 100 及 101 年度財政部中區國稅局某稽徵所轄內遭查獲或輔導補稅之營業人 187 筆樣本及未曾調查補稅之營業人 145 筆樣本，其申報之營業稅及營利事業所得稅資料組合爲 10 個自變數，利用類神經網路及模糊決策樹方法建立選案查核模型，除藉以協助稅務人員挑選可能逃漏營業稅之案件外，並以實際查核情形說明分類矩陣 N01 案件並非單純分類錯誤，反而是稅務人員更應加強注意的案源；另外，比較模糊決策樹可產生語意化規則與類神經網路黑箱作業的特性，藉以說明模糊規則在查核上的實用性，及類神經網路並不適合建立稅務選案模型。

關鍵詞：加值型營業稅、選案查核、類神經網路、模糊決策樹

壹、緒論

賦稅收入爲政府經常支出及公共建設支出最主要的歲入來源，端賴人民的誠實繳納，按照財政部民國 101 年度財政統計年報及財政部財政統計資料庫資料所載，從民國 99 年至 101 年間，政府歲入來源屬稅課收入部分爲民國 99 年度 1,565,847,055 千元、100 年度 1,703,988,545 千元及 101 年度 1,733,349,863 千元，占整體政府整體歲入比例分別爲民國 99 年度 74.02%、100 年度 73.89%及 101 年度 74.67%。而以稅課收入前 3 大稅源營利事業所得稅（以下簡稱：營所稅）、綜合所得稅（以下簡稱：綜所稅）與加值型及非加值型營業稅（以下簡稱：營業稅）比較，於民國 99 年至 101 年間占全國實徵賦稅淨額比例分別爲營所稅 20.34%、綜所稅 20.75%及營業稅 16.68%，顯見該 3 大稅源實屬我國重要歲入之一環。

營業稅性質上爲消費稅（陳清秀，2013），在加值型營業人體系下，再輔以

統一發票的使用，則產生勾稽上、下游營業人交易往來的功能。然而按財政部財政統計資料庫資料，分別比較民國 99 至 101 年間營所稅、綜合所得稅及營業稅所屬違章裁罰金額及違章件數，與同期間占整體稅目裁罰金額比例分別爲營所稅 24.15%、綜合所得稅 17.28%及營業稅 34.52%；裁罰件數比例各爲 3,94%、22.55%及 11.52%，顯見無論在裁罰金額或裁罰件數上，營業稅在我國逃漏情形仍屬嚴重，導致營業稅勾稽與收集營業人進項、銷項交易資料等功能無法發揮外，更連帶加劇營業人營所稅及其業主綜合所得稅的逃漏。

另因爲營業登記條件放寬與工商往來頻繁，營業人登記數量逐年增加，而政府機關落實精簡用人政策的情況下，現行稽徵人員工作量與日俱增，基層人員流動頻繁爲一常態，造成稽徵經驗無法傳承；以本研究對象財政部中區國稅局某稽徵所 101 年度爲例，營業人家數約 9,000 餘家，僅配置營業稅稅務人員 2 人，面對數量龐大的營業人相關稅務申報資料，僅能以有限人力選查 300 餘家營業人進行查核，其查核比率僅約爲 3%，而未加審查的營業人比率則高達 96%以上，導致營業人形成投機心態，降低誠實申報意願。

本研究將以 100 年至 101 年間財政部中區國稅局某稽徵所轄內遭補稅及輔導自動補報繳稅額之加值型營業稅營業人，其實際申報稅務資料數據，透過選取有意義的特徵因子，將該等稅務數據加以建立成資料庫，並以類神經網路及模糊決策樹等資料探勘方法，研究加值型營業稅逃漏稅的預測及歸納各特徵因子與營業人逃漏稅之間的關係及規則，以期營業稅稅務人員於挑選應查核案件上，更加地有效率及有規則可循，藉以降低營業人逃漏營業稅的意願，誠爲本研究主要動機。

貳、文獻探討

一、 加值型營業稅違章樣態：

因加值型營業稅稅額計算是以營業人當期銷項稅額，扣減進項稅額（含前期留抵稅額）後之餘額爲當期應納稅額或留抵稅額，故營業人有短、漏報銷售額或虛報進項稅額，皆能降低應納稅額達到逃漏稅的目的，甚者，可虛增留抵稅額申請退稅導致侵蝕國庫，以下利用「稅務違章案件裁罰金額或倍數參考表」所載有關營業稅法第 51 條及第 52 條漏稅處罰的相關規定，介紹營業人逃漏加值型營業稅樣態：

（一） 短、漏報銷售額樣態：

1. 雖銷售貨物或勞務時依規定開立統一發票，但於填具申報書時短、漏報銷售額導致逃漏營業稅。
2. 銷售貨物或勞務時未依規定開立開立統一發票致短、漏報銷售額導致逃漏營業稅。
3. 銷售貨物或勞務時依統一發票使用辦法第 4 條規定得免開立開立統一發票，但於填具申報書時短、漏報銷售額導致逃漏營業稅。

4. 要求上營業階段營業人以下階段營業人名義跳開統一發票並交付，導致現階段營業人逃漏營業稅（實務上稱：跳開發票）。
5. 重複申報銷售退回或折讓導致逃漏營業稅。

（二） 虛報進項稅額樣態：

1. 雖有實際購買貨物或勞務（實務上稱：有進貨事實者），但按營業稅法第 19 條規定其進項稅額不得扣抵銷項稅額，而營業人卻提出扣抵導致逃漏營業稅。
2. 無實際購買貨物或勞務（實務上稱：無進貨事實者），卻購買不法進項統一發票，虛增進項稅額扣抵銷項稅額導致逃漏營業稅。
3. 申報作廢統一發票所載進項稅額扣抵銷項稅額導致逃漏營業稅。
4. 原申報進項發票，因交易往來產生進貨退出或折讓，但未申報相關進項稅額減少導致逃漏營業稅。
5. 兼營營業人未按「兼營營業人營業稅額計算辦法」調整計算不得扣抵比率，虛增進項稅額扣抵銷項稅額導致逃漏營業稅。
6. 以其他營業人之進項統一發票申報扣抵銷項稅額導致逃漏營業稅。

（三） 其他漏稅樣態：

1. 雖有申報銷售額，但將應稅銷售額申報爲免稅銷售額導致短報銷項稅額逃漏營業稅。
2. 雖有申報銷售額，但將應稅銷售額申報爲零稅率銷售額導致短報銷項稅額逃漏營業稅。

二、 加值型營業稅選案實證文獻

本研究透過回顧相關加值型營業稅選案的論文，以比較各實證採用的研究方法，並發掘出有價值的研究變數，並作爲本研究架構的基礎，相關文獻摘要如表 1。

參、研究方法

一、 對象及樣本

本研究業經財政部中區國稅局某稽徵所單位主管核准，取得該所 100 年至 101 年間轄內受輔導自動補報補繳稅額或遭查獲補徵營業稅稅額 3 萬元以上的加值型營業稅營業人案件，做爲補徵稅額營業人樣本（以下簡稱：遭補稅樣本），計 187 筆，另以調查而未查獲或未曾調查的營業人爲未遭補稅樣本，計 145 筆，合計爲 332 筆，樣本範例如表 2。同時，以樣本營業人當年度申報營利事業所得稅主要行業代號爲主，按稅務行業代號標準分類區分爲製造業、買賣業、服務業及營造業等 4 大類，作爲模糊決策樹學習組及測試組的資料來源。

表 1 偵測加值型營業稅逃漏稅有關文獻彙總表

研究者	研究題目	採用研究方法	研究變數
陳國元（2002）	加值型營業稅中營業人涉嫌虛設行號之實證研究	區別分析 邏輯斯迴歸	進項金額等 12 個變數
陳彥文（2003）	類神經網路之應用-以營業稅選案模型爲例	類神經網路	本期應納稅額÷銷售額總計等 5 個變數
江振豐（2006）	營業稅選案查核-以類神經網路建構選案模型	類神經網路 邏輯斯迴歸	t 期加值率等 11 個變數
林惠英（2006）	應用資料探勘技術於營業稅逃漏稅選案之研究	關聯式法則	稅籍主檔等資料
邱家永（2008）	加值型營業稅冒退稅選案查核模式之研究	區別分析 邏輯斯迴歸	零稅率銷售額÷銷售總額等 11 個變數
林月桃（2009）	影響逃漏稅之因素-以加值型營業稅爲例	邏輯斯迴歸	稅籍變更頻繁等 12 個變數
謝碧霞（2010）	虛設行號查緝之研究-以財政部臺灣省中區國稅局查核案件爲例	區別分析 邏輯斯迴歸	營業人申報加值率等 4 個變數
胡萬能（2011）	運用決策樹演算法於營業稅選案之研究	決策樹	營業稅平均銷售額等 25 個變數
王淑芳（2010）	營業人加值型營業稅逃漏稅金額之決定因素-以內湖科學園區逃漏稅營業人爲例	迴歸分析	違章年度銷售金額等 23 個變數
洪梓玲（2011）	建構加值型營業稅選案查核模式之研究	決策樹 區別分析 類神經網路	營業稅申報加值率等 5 個變數
張春蓮（2013）	加值型營業稅逃漏偵測之研究	決策樹 邏輯斯迴歸 區別分析	營業稅申報加值率等 21 個變數

資料來源：本研究整理

表 2 營業人申報資料樣本範例

年度	統編	名稱	主要行業代號	行代類別
101	27○○○○○○	○○公司	461517	2（買賣業）
同業標準毛利率	當年遭補稅	當年銷售額	當年進項金額	銷貨收入淨額
18%	2○○, ○○6	37,638,754	46,098,357	37,638,754
當年毛利率	當年費用率	當年純益率		
16.69%	14.02%	2.07%		
以下爲年底資產負債表科目	現金	銀行存款	流動金融資產	應收票據
	47,030	1,852,406	0	4,533,300
應收帳款	其他應收款	存貨	預付款	其他流動資產
382,955	24	48,584,359	37,341	752,929
短期借款	應付款項	預收款項	其他流動負債	長期負債
23,527,925	1,970,597	0	22,724,194	0
其他負債	資產總額			
51,500	57,145,637			
100 年存貨				
35,422,094				

資料來源：財政部中區國稅局稅務系統

二、 研究變數

本研究設計的應變數及自變數，除參考文獻及財務報表分析（謝劍平，2012）外，另參酌財政部中區國稅局內部營業稅資深人員訪談建議，而該等變數的定義如表 3。

表 3 各變數定義

變數	符號	變數	符號
$\frac{\text{當年度申報毛利率}}{\text{當年度同業利潤標準毛利率}}$	X1	$\frac{\text{當年度申報加值率平均值}}{\text{當年度同業利潤標準毛利率}}$	X2
$\frac{\text{當年底存貨數}}{\text{當年度申報銷售額}}$	X3	$\frac{\text{當年底其他流動負債總額}}{\text{當年度申報銷售額}}$	X4
$\frac{\text{當年度成本及費用}}{\text{當年度申報進項發票金額}}$	X5	$\frac{\text{當年度進貨金額}}{\text{當年度申報進項發票金額}}$	X6
行業別	X7	$\frac{1}{\text{流動比率}}$	X8
$\frac{\text{當年底負債總額}}{\text{當年底總資產}}$	X9	$\frac{\text{當年度申報銷售額}}{\text{當年底總資產}}$	X10

資料來源：本研究整理

三、 類神經網路

按 Berry and Linoff （1997）類神經網路的基本單位稱爲處理單元(processing unit)，如圖 1，其概念來自於模擬生物神經單元的運作方式，處理單元可接受多個輸入，並以數學函數加以處理並產生單一輸出。處理單元內的數學函數可分成 2 部份：組合函數及轉換函數。組合函數有很多種，最常見的就是將輸入乘上權重後的加總，而轉換函數則負責將組合函數的輸出轉換成單元的輸出，最常使用的轉換函數爲 Sigmoid 函數，使其輸出界於 0 與 1 之間，當輸入極小時，輸出爲 0，當輸入極大時，輸出爲 1。

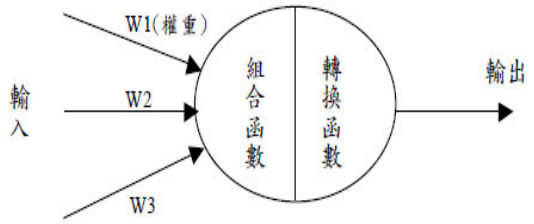


圖 1 處理單元之基本構造

類神經網路可分爲 3 層：輸入層（Input Layer）、隱藏層（Hidden Layer）及輸出層（Output Layer），如圖 2。輸入層接受外界輸入，其處理單元數即爲輸入變數的數量，隱藏層，爲類神經網路中最重要的部份，通常爲 1 至 2 層，而輸出層，爲輸出網路的處理結果，輸出層的處理單元數爲研究問題所欲得到結果的數量。

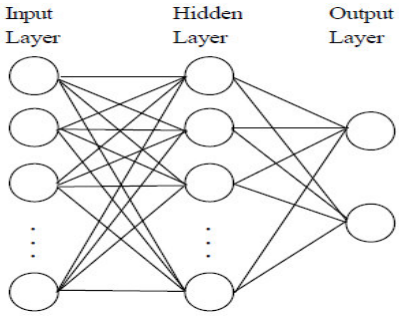


圖 2 類神經網路基本架構

四、 模糊決策樹

在大部分決策樹歸納方法常被使用在分類問題，並建立專家系統的知識，但其分類規則卻無法以人類認知的方式來表達及處理問題，故有學者對於輸出變數屬於類別型的問題，將 Quinlan(1986)提出原本用來分類類別型資料的 ID3(Iterative Dichotmiser 3)決策樹演算法，以 Zadeh(1965)提出的模糊理論改良爲模糊決策樹(Fuzzy Decision Tree)；前後者間最主要差異在於 ID3 演算法是以各變數對應於樣本的不確定性-熵值(entropy)及各變數對

於樣本的資訊增益(information gain)作爲節點分支準則，而 Weber(1992)則以歸屬程度(membership degree)，或者有 Yuan and Shaw(1995)以分類模糊度(classification ambiguity)等模糊理論觀念，改寫 ID3 演算法熵值的計算公式，並作爲模糊決策樹的分支準則。以下予以簡述模糊決策樹重要觀念：

（一） 模糊理論：

在現實生活中，常常會有許多情況和感受是難以精確描，且無法以量化的方式表達，如優劣、大小或尚可等皆屬於人類使用的口語，其無法以一個明確的界線來區分，這就是模糊。而模糊具體而言是指客觀事務的明顯差異間，有其過渡時所呈現的「亦此亦彼」，美國加州大學柏克萊分校的 Zedeh 教授於 1965 年建構出模糊邏輯及模糊測量理論，Zedeh 利用隸屬程度來描述差異的中間過度，是用精確的數學語言對模糊性的描述(萬絢、林明毅等 2006)。

（二） 歸屬函數：

模糊集合是指用來表示界線或邊界不明確且具有特定性質的事物集合，其觀念則是說 x 是否屬於集合 A 不再是一個若非即是的問題，而用 0 至 1 的歸屬函數值來表示 x 屬於 A 的程度。舉例三角形歸屬函數圖形如圖 3，其定義如公式(1)：

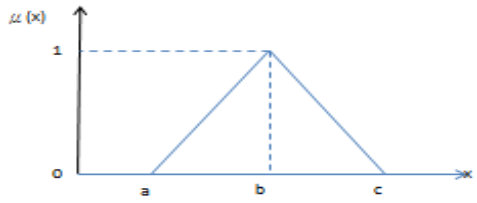


圖 3 三角形歸屬函數

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & if \ \ a \le x < b \\ 1, & if \ \ x = b \\ \frac{c-x}{c-b}, & if \ \ b < x \le c \\ 0, & if \ \ x < a \ \ or \ \ x > c \end{cases} \qquad (1)$$

（三） 語意變數：

人類思維與交際使用的語言稱爲自然語言，自然語言的一個重要特徵就是具有模糊性。語意變數(linguistic variables)就是用人類自然語言中的語詞來表示對某人事物的感知程度，而非以數值型態來傳達，例如我們常用的語言「很好」、「好」、「普通」、「差」、「很差」等，特點就在於其模糊性，所以稱之爲「模糊詞」。如果將欲形容事物(變數)範圍內的模糊詞用「或」、「且」等連接起來，就可以利用隸屬函數來進行數學化(萬絢、林明毅等 2006)。

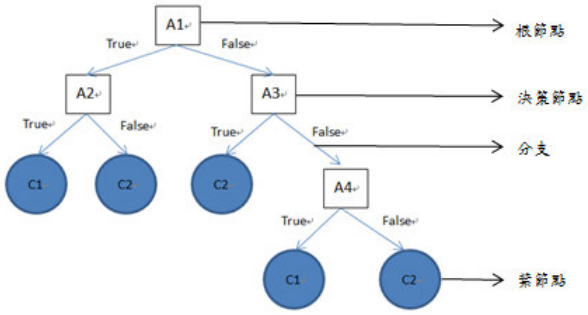
（四） 解模糊化：

解模糊化(Defuzzication)是將模糊集合資料轉換爲一明確的資料。解模糊化的方法有許多不同方式，可視問題特性而選定使用，一般常用的

方式爲 Takagi and Sugeno(1985)的平均權重法。

（五） 決策樹：

決策樹(Decision Tree)是一個重要的資料探勘方法，其代表著決策集合的樹形結構，爲一種監督式學習的方式，其基本組成部分有：決策節點、分支和葉節點。決策樹中最上面的節點稱爲根節點(root)，是整個決策樹的開始。每個分支是一個新的決策節點，或者是樹的葉節點，每一個決策節點代表一個問題或決策，通常對應於等待分類物件中的屬性。每一個葉節點代表一種可能的分類結果。在沿著決策樹從上到下的過程中，在每個節點都會遇到一個測試，對每個節點上不同問題的測試結果將會導致不同的分支，最後會到達一個葉節點。其基本示意圖如圖 4：



An 爲輸入變數屬性；C1 及 C2 爲分類結果

圖 4 決策樹示意圖

五、 評估方法

本研究將過去文獻使用之分類矩陣修改爲適合稅務人員閱讀的稅務選案矩陣，如表 4，主要是對 N01 分類以「未逃漏稅分類爲逃漏稅」表達時，容易讓使用人單純誤解爲模型分類錯誤，故修改爲「未遭補稅樣本分類爲應查核案件」後，方能明確表達出 N01 案件實際是稅務人員應投入更多心力查核的案源。

表 4 稅務分類矩陣

		預測分類		合計
		應查核案件	無須查核案件	
輸入樣本	遭補稅樣本	N11	N10	N3
	未遭補稅樣本	N01	N00	N4
合計		N1	N2	N

資料來源：本研究整理

肆、研究結果

一、 類神經網路-未分業別樣本

本研究按 Lippmann(1987)、Hecht-Nielsen(1990)及葉怡成（2009）對於隱藏層的處理單元數量設定建議，以 2 組網路架構「10-6-2」及「10-21-2」，設定學習速率在 0.1 至 1.0，間距 0.1，動量在 0 至 0.9，間距 0.1 及學習次數 1,000 次與 2,000 次，以全部樣本取 60%爲學習組與 40%爲測試組，並記錄實證分類結果計 400 次，其中以網路設定 10-21-2、學習速率 1.0、動量 0.9、學習次數 2,000 次，該組設定整體正確率最佳爲 65.41%，其分類稅務分類矩陣如表 5；而該組設定下各變數的敏感度分析如表 6。

表 5 正確率最佳組別稅務分類矩陣

		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	53	17	70
	未遭補稅樣本	29	34	63
合計		82	51	133

資料來源：本研究整理

表 6 正確率最佳組別自變數敏感度分析（單位：%）

變數	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
對 Y 敏感度	5.63	16.64	6.12	18.11	9.40	9.95	5.73	3.02	4.54	20.87

資料來源：本研究整理

二、 模糊決策樹-未分業別樣本

（一） 變數模糊化

本研究利用 Guillaume et. al.（2011）撰寫之 FisPro 軟體，先對全部樣本 332 組樣本的 60%作爲學習組，40%作爲測試組，將各自變數的歸屬函數以平均方式區分 5 個模糊分群，並加以語意化，列舉變數 X1 及 X7 如表 7。

表 7 未分業別變數模糊集合及語意程度表

變數 X1					變數 X7		
MF1	-0.767	-0.767	0.214	低	MF1	1	製造業
MF2	-0.767	0.214	1.196	中低	MF2	2	買賣業
MF3	0.214	1.196	2.177	中	MF3	3	服務業
MF4	1.196	2.177	3.158	中高	MF4	4	營造業
MF5	2.177	3.158	3.158	高			

資料來源：本研究整理

（二） 產出模糊規則

以學習組爲監督學習樣本，產出模糊規則，如表 8。

表 8 未分業別學習組模糊規則

規則	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
1		中		正常	正常		製造				未漏
2		中		正常	正常		營造				未漏
3		中		異常							未漏
4		中高									未漏
5		高				尚可					未漏
6		高				高					未漏
7		高				異常					未漏
8		低									逃漏
9		中低									逃漏
10		中		正常	正常		買賣				逃漏
11		中		正常	正常		服務				逃漏
12		中		正常	尚可						逃漏
13		中		正常	高						逃漏
14		中		正常	異常						逃漏
15		中		尚可							逃漏
16		中		高							逃漏
17		中		過高							逃漏
18		高									逃漏

資料來源：本研究整理

（三） 測試組分類結果

觀察未分業別的規則僅受 X2、X4、X5 及 X7 等變數影響，而變數 X1、X3、X6 及 X8 至 X10 對於營業人逃漏稅無明顯影響，各規則間無明顯差異及特徵，該等規則是否合理，資深稅務人員難以判斷並據以修改，故逕以該等規則對測試組樣本做分類，如表 9。

表 9 未分業別測試組稅務分類矩陣

		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	65	10	75
	未遭補稅樣本	48	10	58
合計		113	20	133

資料來源：本研究整理

三、 模糊決策樹-分業別樣本

（一） 變數模糊化

本研究將全部樣本按行業別區分爲 4 組樣本，分別爲製造業 30 組、買賣業 224 組、服務業 48 組及營造業 30 組，同樣以各組樣本的 60%作爲學習組，40%作爲測試組，另剔除變數 X7，其餘過程與上述肆、二、（一）相同。

（二） 產出模糊規則

以各業別之學習組爲監督學習樣本，產出模糊規則，並針對各業別特性及資深稅務人員意見修該規則，原規則及修改後規則如表 10，對測試組分類結果如表 11，修改理由說明如下：

1. 製造業規則：分類結果不盡理想，重新檢視原規則；比較規則 4、5、6、8 及 9，可觀察出規則 4 與規則 9，在 X1 及 X2 皆爲相同的特徵程度下，其 X10 偏低的規則 4 爲對應未逃漏稅 Y，而中等的規則 9 卻對應逃漏稅 Y，顯有不合理處，經修正後規則後，重新對測試組樣本分類，可觀察到遭補稅樣本分類正確性大幅上升，而未遭補稅樣本僅多 1 樣本遭分類至應查核案件。
2. 買賣業規則：觀察規則 5、16 及 17，規則 16 及 17 的變數 X8 尚屬合理，但規則 5 在流動比率仍屬低點的情形下，對應 Y 未漏稅，經資深稅務人員建議，不予採用規則 5；另規則 7 的流動比率屬異常情形，亦對應 Y 未漏稅，相同理由不採用，經修正規則後，重新對測試組樣本分類，可觀察到遭補稅樣本分類正確性再上升，而未遭補稅樣本分類情形不變。
3. 服務業規則：分類結果不盡理想，分析服務業因爲勞力密集原因，通常營業稅加值率都會偏高，但營業成本會有直接人工關係，毛利率不一定會偏高，比較規則 4、5、13 及 14，同樣在 X1 偏低情形下，X6（當年進貨進額÷當年申報進項發票金額，依查核經驗該比率偏高時，應注意其有取具進項發票，但以未提出申報扣抵銷項稅額方式框飾營業稅申報書。）偏高的規則對應 Y 逃漏稅，尚屬合理。另服務業不同於買賣業的特徵在於其他負債科目，服務業別如

美容業通常有預收款項，該部分雖尚未認列銷貨收入，但營業人應於收款時開立發票申報銷售額，如漏報該部分銷售額時，連帶會影響其他負債科目偏低，故比較規則 1 及規則 10，其在 X4 部分可能有不合理之處，按資深稅務人員建議將該部分改寫，經修正規則後，重新對測試組樣本分類，可觀察到遭補稅樣本分類正確性大幅上升，而未遭補稅樣本分類出 4 件應查核案件。

4. 營造業規則：產出規則有 19 條，但本研究營造業學習組的輸入樣本僅 18 個，顯見營造業學習組樣本間並無共同營業稅逃漏稅的規則，只能對學習組樣本做個別規則解釋，各不再做修正；而該模糊規則對測試組的分類結果亦不盡理想，但實務上營造業的銷貨收入及成本費用多是以完工比率法認列，但其營業稅銷售額則是依據「營業人開立銷售憑證實現表」包作業規定，依其工程合約所載每期應收價款時開立統一發票開立及申報，故營業稅的銷售額與營所稅的銷貨收入通常會產生跨年度的落差，而本研究選擇的變數是以營業人當年的營所稅及營業稅申報資料組合而成，應該是該模糊規則對測試組分類不佳的主要原因。

表 10 分業別學習組模糊規則

製造業模糊規則										
規則	X1	X2	X3	X4	X5	X6	X8	X9	X10	Y
1	中低									未漏
2	中	低								未漏
3	中	中低								未漏
4	中	中							中低	未漏
4*	中	中							中	未漏
5	中	中							中高	未漏
6	中	中							高	未漏
7	低									逃漏
8	中	中							低	逃漏
9	中	中							中	逃漏
9*	中	中							中低	逃漏
10	中	中高								逃漏
11	中	高								逃漏
12	中高									逃漏
13	高									逃漏
買賣業模糊規則										
規則	X1	X2	X3	X4	X5	X6	X8	X9	X10	Y
1	中	中		正常			尚可	尚可	中低	未漏

2	中	中		尚可			尚可	尚可	中	未漏
3	中	中		尚可			尚可	尚可	中高	未漏
4	中	中		尚可			尚可	尚可	高	未漏
5	中	中		尚可			低			未漏
5*	不採用									
6		中高		尚可			尚可			未漏
7		中高					異常			未漏
7*	不採用									
8		高								未漏
9		低								逃漏
10		中低								逃漏
11	低	中								逃漏
12	中低	中								逃漏
13	中	中		尚可			尚可	尚可	低	逃漏
14	中	中		尚可			尚可	高		逃漏
15	中	中		尚可			尚可	過高		逃漏
16	中	中		尚可			過低			逃漏
17	中	中		尚可			異常			逃漏
18	中	中		高						逃漏
19	中	中		過高						逃漏
20	中高	中								逃漏
21	高	中								逃漏
22		中高		高			尚可			逃漏
23		中高		過高			尚可			逃漏
24		中高					低			逃漏
25		中高					過低			逃漏
26		中高					異常			逃漏
服務業模糊規則										
規則	X1	X2	X3	X4	X5	X6	X8	X9	X10	Y
1		中		低			尚可			未漏
1*		中		中低			尚可			未漏
2		中					過低			未漏
3	低	中高								未漏
4	中低	中高				可				未漏
5	中低	中高				尚可				未漏
6	中	中高								未漏
7		高							高	未漏

8		低								逃漏
9		中低								逃漏
10		中		中低			尚可			逃漏
10*		中		低			尚可			逃漏
11		中					異常			逃漏
12		中					異常			逃漏
13	中低	中高				中				逃漏
14	中低	中高				過高				逃漏
15	中高	中高								逃漏
16	高	中高								逃漏
17		高							低	逃漏
18		高							中低	逃漏
營造業模糊規則										
規則	X1	X2	X3	X4	X5	X6	X8	X9	X10	Y
1		中低						低		未漏
2	中高	中		過低				低	低	未漏
3		中		過低				低	中低	未漏
4		中高							低	未漏
5		中高							中低	未漏
6	高	中高							中	未漏
7		高								未漏
8		低								逃漏
9		中低						中低		逃漏
10		中低						中		逃漏
11		中低						中高		逃漏
12	中	中		過低				低	低	逃漏
13		中		低				低		逃漏
14		中						中		逃漏
15		中						中高		逃漏
16	中	中高							中	逃漏
17	中高	中高							中	逃漏
18		中高							中高	逃漏
19		中高							高	逃漏

規則旁加註*者爲修改後規則

資料來源：本研究整理

表 11 分業別測試組稅務分類矩陣

製造業		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	5（1）	1（5）	6
	未遭補稅樣本	3（2）	3（4）	6
合計		8（3）	4（9）	12
買賣業		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	46（44）	6（8）	52
	未遭補稅樣本	22（22）	16（16）	38
合計		68（66）	22（24）	90
服務業		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	11（6）	1（6）	12
	未遭補稅樣本	4（0）	4（8）	8
合計		15（6）	5（14）	20
營造業		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	3	2	5
	未遭補稅樣本	3	4	7
合計		6	6	12
合併		預測分類		合計
		應查核案件	無須查核案件	
輸入	遭補稅額樣本	65（54）	10（21）	75
	未遭補稅樣本	32（27）	27（32）	59
合計		97（81）	37（53）	134

表（ ）內爲原規則對測試組分類數據

資料來源：本研究整理

四、 分類模型比較

比較未分業別類神經網路與未分業別模糊決策樹分類矩陣，再以分業別模糊決策樹分類矩陣比較前開 2 者，如表 12，茲分析說明如下：

表 12 各分類模型比較

		未分業別類神經網路		未分業別模糊決策樹		分業別模糊決策樹	
		應查核案件	無須查核案件	應查核案件	無須查核案件	應查核案件	無須查核案件
輸入樣本	遭補稅樣本	53	17	65	10	65	10
	未遭補稅樣本	29	34	48	10	32	27
型 1 正確率		75.714%		86.667%		86.667%	
型 2 正確率		53.968%		17.241%		45.763%	
整體正確率		65.414%		56.391%		68.657%	

資料來源：本研究整理

（一） 未分類樣本分類比較：

本研究在未分類樣本條件下,分類模型的型 1 正確率以模糊決策樹(以下稱：模型 2）的 86.667%較高，因爲本研究的主要目的是預測營業人是否逃漏營業稅，故因以型 1 較高的分類模型爲首選，但在比較型 2 正確率時，則可觀察到類神經網路模型（以下稱：模型 1）及模型 2 的型 2 正確率不但較型 1 爲低，且有正確率偏低情形，其主要原因在於實際稽徵業務上未遭調查的營業人數目上遠大於遭調查補稅的營業人，故研究樣本所挑選的「未漏稅」樣本，實際上應爲「受調查而未查獲或未曾調查的營業人」樣本，而參雜了可能逃漏稅的樣本，導致各分類模型的型 2 正確率偏低；若以稅務人員角度而言，如將分類矩陣內的 N01 案件直接視爲模型的「分類錯誤」，而逕行討論何種分類方法的型 2 正確率較佳，似欠妥當，因爲該等案件應該是稅務人員感到有興趣查核的案件。

（二） 未遭補稅樣本分類爲應查核案件（N01 案件）的解決方法：

N01 的案件，可視爲分類模型爲稅務人員挑選出應查核案件，但必須釐清的一個稽徵實務事實是選出的案件不能未經查核就逕行做核定處分，故稅務人員必須要確實瞭解模型爲何分類出 N01 案件。以類神經網路而言，本研究分析了各自變數對是否逃漏稅的敏感度後，可觀察到其中變數 X2、X4 及 X10 對營業人是否逃漏稅「各自」極具影響，但實務上資深稅務人員判斷營業人是否漏稅的思考過程是各種情況（即各變數或特徵因子）的綜合考量，故類神經網路分類過程的黑箱作業，對於缺乏經驗的稅務人員而言，即無法理解的爲何類神經網路模型將未遭補稅樣本分類爲應查核案件的理由，反映於實際查核 N01 案件上，就難免迷失調查方向而徒增稽徵成本；另未分業別下的模糊決策樹模型，其產出的規則僅由 X2、X4、X5 及 X7 等變數影響，而變數 X1、X3、X6 及 X8 至 X10 對於逃漏稅卻無影響，各規則間無明顯差異及特徵，未能符合資深稅務人員經驗，致難以判斷該等規則是

否合理，故本研究針對查核實務所需，另以分業別樣本方式以模糊決策樹建立分類模型（以下稱：模型 3）。

（三） 以分業別模糊規則確立 N01 案件調查方向：

觀察模型 3 的型 1 正確率同樣達到 86.667%，但其 N01 分類案件卻不至於如模型 2 般，將大部分未漏稅樣本分類爲應查核案件，而徒增稽徵成本，另因爲模型 3 產出的模糊規則，除能表現出各業別可能逃漏稅的特徵外，其具有語意化的特性，符合人類的思考模式而具有可讀性，故能加以解釋及做合理修改，茲分析修改後規則如下：

1. 修改後製造業規則：製造業的毛利率（X1）及加值率（X2）過度偏離同業利潤標準毛利率時，縱然是偏高，營業人仍然有逃漏稅可能；而毛利率及加值率未偏離同業利潤標準毛利率時，配合觀察銷售規模（X10），其偏低時營業人因營業風險增加，致逃漏稅可能性愈大。
2. 修改後買賣業規則：買賣業的毛利率及加值率與同業利潤標準毛利率比較下，如是持平或偏高者，營業人逃漏稅可能性較小，但相同條件下，如果營業人有其他負債（X4）、流動比率偏低（X8）、負債比率偏高（X9）或銷售規模（X10）偏低等增加財務風險的情形下，其逃漏稅可能性增加；另加值率（X2）偏高，但伴隨著其他負債偏高或流動比率偏低，營業人以漏進漏銷方式框飾營業稅申報書的可能性增加。
3. 修改後服務業規則：服務業因爲勞力密集因素（執發扣繳憑單，而非取得進項發票），多有加值率（X2）偏高情況；本研究按資深查核人員建議將規則 1 及規則 10 的其他負債（X4）修正後，大幅改善的分類結果，證明了服務業的其他負債不應過低，明顯與買賣業不同；另營業成本（X6）偏高時，營業人可能隱匿銷售勞務收入（X6：當年度進貨金額÷當年度申報進項發票金額，同服務業的年底存貨數較固定，故可將當年度進貨金額簡化視爲當年度營業成本，故買賣業的營業成本內涵發票憑證偏低時，可推論其勞務收入比率較大）；而流動比率（X8）及銷售規模（X10）的判斷上如同買賣業。
4. 營造業因規則數大於樣本數，無樣本間共同逃漏稅特徵，故不做分析。

（四） 比較結論：

因爲分業別的模糊規則具有可讀性，查核人員對於不孰悉的案件（即分類矩陣內 N01 的案件），可運用歷史案件所建立的規則，在各變數的相互變動間尋找出相似處，如同複製了一位資深稅務人員的經驗法則，俾便迅速確立調查方向，因此對於分類矩陣的 N01，本研究認爲應利用分業別的模糊規則來解釋其爲何應予查核及確立調查方向，相

反地，也能解釋分類爲 N00 的案件爲何無須調查。

伍、結論

一個解決問題的模型應確實考慮實際使用者的情況，當解釋能力很重要，使用者必須知道輸入變數對輸出變數的重要程度及關係時，許多使用者常採用較具解釋能力的分析工具，而捨棄具有建構準確非線性模型能力的分析工具(葉怡成，2009)。本研究的研究目的在於幫助查核人員有效的挑選出可能逃漏加值型營業稅的營業人，本研究建立的 3 個分類模型互相比較後，分業別模糊決策樹模型的型 1 正確率 86.667%，除能有效將逃漏稅樣本分類爲應查核案件外，相較於類神經網路輸出、入的黑箱對應或未分業別模糊規則的無法閱讀，其產出的分業別模糊規則，因具有可讀性，能有效幫助查核人員解決分類矩陣中 N11 及 N01 的查核案件。

參考文獻

1. 王淑芳（2010）。營業人加值型營業稅逃漏金額之決定因素-以內湖科學園區逃漏稅營業人爲例（未出版之碩士論文）。國立政治大學，台北市。
2. 加值型及非加值型營業稅法(2014)。
3. 江振豐（2006）。營業稅選案查核-以類神經網路建構選案模型。雲林科技大學，雲林縣。
4. 邱家永（2008）。加值型營業稅冒退稅選案查核模式之研究。國立臺中技術學院，台中市。
5. 林惠英（2006）。應用資料探勘技術於營業稅逃漏稅選案之研究。國立中正大學，嘉義縣。
6. 林月桃（2009）。影響逃漏稅之因素-以加值型營業稅爲例。國立政治大學，台北市。
7. 胡萬能（2011）。運用決策樹演算法於營業稅選案之研究。元智大學，桃園縣。
8. 洪梓玲（2011）。建構加值型營業稅選案查核模式之研究。臺中技術學院，台中市。
9. 財政部（2012）。民國 101 財政統計年報。台北市：財政部統計處。
10. 財政部. 財政部財政統計資料庫。取自：http://web02.mof.gov.tw/njswww/WebProxy.aspx?sys=100&funid=defjspf2
11. 財政部. (2012)。中華民國稅務行業標準分類(第 7 次修訂)。台北市：財政部統計處。
12. 兼營營業人營業稅額計算辦法(2006)。

13. 陳國元（2002）。加值型營業稅中營業人涉嫌虛設行號之實證研究。朝陽科技大學，台中市。
14. 陳彥文（2003）。類神經網路之應用－以營業稅選案模型爲例。銘傳大學，台北市。
15. 陳清秀(2012)。營業稅之稅捐主體與客體之探討。當代財政，25，19-40。
16. 張春蓮（2013）。加值型營業稅逃漏偵測之研究。國立臺中科技大學，台中市。
17. 葉怡成（2009）。類神經網路模式應用與實作。台北市：儒林圖書有限公司。
18. 稅務違章案件裁罰金額或倍數參考表（2013）。
19. 萬絢、林明毅、陳宏杰(2006)。模糊理論應用與實務。台北市：儒林圖書有限公司。
20. 謝碧霞（2010）。虛設行號查緝之研究-以財政部臺灣省中區國稅局查核案件爲例。台中市，臺中技術學院。
21. 謝劍平（2012）。財務報表分析。台北市：智勝文化事業有限公司。
22. Berry, M. J., & Linoff, G. (1997). *Data mining techniques: for marketing, sales, and customer support*. John Wiley & Sons, Inc..
23. Guillaume, S., & Charnomordic, B. (2011). Learning interpretable fuzzy inference systems with FisPro. *Information Sciences*, 181(20), 4409-4427.
24. Hecht-Nielsen, R. *Neurocomputing*, 1990. Addison-Wisley, Sydney.
25. Lippmann, R. P. (1987). An introduction to computing with neural nets. *ASSP Magazine, IEEE*, 4(2), 4-22.
26. Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.
27. Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, (1), 116-132.
28. Weber, R. (1992, July). Fuzzy-ID3: a class of methods for automatic knowledge acquisition. *In The second international conference on fuzzy logic and neural networks* (pp. 265-268).
29. Yuan, Y., & Shaw, M. J. (1995). Induction of fuzzy decision trees. *Fuzzy Sets and systems*, 69(2), 125-139.
30. Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.