

地方稅務系統應用模糊理論進行房屋稅資料探勘之實務研究

王智昌、晁瑞明

E-mail: 9314336@mail.dyu.edu.tw

摘要

本研究主要以我國現有地方稅務系統的房屋稅稅籍資料中的稅籍主檔為對象，使用Apriori演算法配合模糊理論，進行房屋稅資料探勘，來找尋區域建築物「構造別、使用別」間的關聯性規則。施政者在進行區域建設規畫時，往往缺少強力的數據為佐證，對該區域進行最有利建設，而只是規畫者隨著主觀的意念進行建設，為解決上述的問題，都市建設規畫者能透過本系統，可找出個別區域中建築物之構造、使用類別的關聯性，於都市計畫區域開發設計時，能搭配區域建築物特性及使用類別，做出符合區域最佳的整體規畫及建設。從彰化縣一百五十七萬筆房屋稅紀錄中，統計出各鄉鎮市每個里（村）中的二十四種「構造別、使用別」所持有的比例，發現愈都市化的區域，建物構造別以混凝土居多，鄉下則以磚造居多、木土造居次，與一般所持觀點吻合，本研究可提供確切數據證明此一觀點。居住區域的建物構造特色及使用別的關聯性，關係著公共安全及居住品質，針對統計出來的每個里（村）的「構造別、使用別」資料，選出目標區域中（包含數個里）進行關聯性探勘，找出如「混凝土、自住」與「磚造、自住」有關聯性，顯示國人會於房屋完工後，在頂樓加蓋違章建築的習性有關，主政者可利用此一關聯性進行如拆違建等行政政策或修改法令。

關鍵詞：房屋稅、Apriori、模糊關聯性、資料探勘

目錄

目錄封面內頁 簽名頁 授權書	iii 中文摘要	v 英文摘要
vi 誌謝	vii 目錄	
viii 圖目錄	xi 表目錄	
xiii 第一章 緒論 1.1 研究背景	1 1.2 研究動機	3
1.3 研究目的	4 1.4 研究對象及限制	5 1.5 研究方法
5 1.6 研究架構	7 第二章 文獻探討 2.1 地方稅務系統現況	8 2.1.1
地方稅務系統架構	8 2.1.2 地方稅務系統稅務資料庫簡述	10 2.2 資料
庫	11 2.2.1 資料庫系統的組成	12 2.2.2 資料庫的分類
查詢語言	13 2.3 資料探勘	14 2.3.1 資料探勘概述
探勘的功能	15 2.3.3 資料探勘的方式	19 2.3.4 Apriori演算法
的四個階段	23 2.4 模糊理論	25 2.4.1 模糊理論的基本概念
26 2.4.3 語意變數	27 2.4.4 模糊集合的運算	27 2.4.5 解模糊
2.4.6 模糊查詢系統	28 第三章 研究方法 3.1 建立SQL Server資料庫	30 3.2 資料轉換
32 3.3 定義需解決的問題	32 3.4 資料預處理	33 3.5 語義模糊化
34 3.6 以Apriori演算法進行資料探勘	36 3.7 結果解模糊	38 3.8 系統流程圖
41 第四章 系統建立 4.1 系統建構環境與開發工具	40 4.2 房屋稅資料分析與轉換	41 4.3 系統程式流程與
功能說明	45 4.3.1 里(村)構造及使用別查詢	47 4.3.2 區域關聯性查詢
5.1 「構造、使用別」面積的統計結果	56 5.2 關聯性資料探勘的結果	60 第六章 結論 6.1 結論
研究	65 參考文獻	66 圖目錄 圖1.1 研究架構圖
9 圖2.2 高頻項目集的生成過程	22 圖3.1 系統架構圖	31 圖3.2 比例的隸屬函數
統流程圖	39 圖4.1 房屋稅課稅主檔格式	41 圖4.2 挑錄後的房屋稅資料紀錄
格式	43 圖4.4 鄉鎮市村里中文對照檔	44 圖4.5 系統流程圖
圖4.7 s0.asp網頁	48 圖4.8 s1.asp網頁	48 圖4.9 s2.asp網頁
50 圖4.11 s4.asp網頁	51 圖4.12 s5.asp網頁	51 圖4.13 d1.asp網頁
圖4.15 d3.asp網頁	54 圖4.16 d4.asp網頁	54 圖4.17 d5.asp網頁
56 圖5.2 鄉鎮市面積比統計之二	57 圖5.3 鄉鎮市面積比統計之三	57 圖5.4 鄉鎮市構造自住比例分析
圖 59 圖5.5 和美鎮地圖	60 圖5.6 和美鎮資料探勘結果顯示	61 圖5.7 溪州鄉關聯性探勘結果輸入
圖5.8 溪州鄉關聯性探勘結果顯示	62 表目錄 表2.1 Data Mining運用的理論與實際應用功能比較表	18 表2.2 每戶房屋面
積比例資料庫A	21 表3.1 面積比統計表	35 表3.2 模糊集合表
35 表3.3 C1模糊值計算表	37 表4.1 房屋	37 表4.1 房屋
違建類型及面積比例	55 表5.1 台閩地區使用執照核發面積一覽表	58 表6.1 鄉鎮市「構造、使用別」統計表
64		

參考文獻

1. 王文俊(2001), 認識FUZZY (第二版), 全華科技圖書股份有限公司。
2. 王瑞豐(2001), 「影響地方稅稽徵機關導入流程再造之關鍵因素探討」, 中正大學資訊管理學系碩士論文。
3. 王景聰(2003), 「運用資料挖掘技術於信用卡顧客關係管理之研究」, 元智大學工業工程與管理研究所碩士論文。
4. 尹相志(2003), SQL 2000 Analysis Service資料探勘服務, 維科圖書有限公司。
5. 中華資料採礦協會(2004), 「Data Mining 方法介紹」, 網頁 (<http://www.cdms.org.tw/index.html>), 中華資料採礦協會。
6. 內政部營建署(2004), 「臺閩地區核發建築執照統計」, 網頁 (<http://www.cpami.gov.tw/pws/pws1.php>), 內政部營建署。
7. 行政院研究發展考核委員會(2002), 「電子化政府報告書」, 行政院研究發展考核委員會。
8. 位元文化(2002), ASP 3.0動態網頁入門實務, 文魁資訊股份有限公司。
9. 李卓翰(2002), 資料倉儲 (理論與實務), 學貫行銷股份有限公司。
10. 李允中、王小璠、蘇木春(2002), 模糊理論及其應用, 全華科技圖書股份有限公司。
11. 吳琮璠、謝清佳(2002), 資訊管理理論與實務, 智勝文化事業有限公司。
12. 林騰蛟、曹祥雲(2002), 系統分析與設計 Systems Analysis and Design, 新文京開發出版有限公司。
13. 邱永祥 (2003), 「運用類神經網與資料探勘技術於網路教學課程推薦之研究」, 朝陽科技大學資訊管理學系碩士論文。
14. 法務部全國法規資料庫(2004), 「都市計畫法」, 法務部全國法規資料庫網站(<http://law.moj.gov.tw>), 法務部。
15. 秉昱科技編譯(2001), 模糊邏輯與類神經模糊在商業和財政的應用, 儒林圖書有限公司。
16. 施威銘研究室(2002), Microsoft SQL Server 2000設計實務, 旗標出版股份有限公司。
17. 財政部財稅資料中心 (1999), 財稅資訊作業規範 - 房屋稅 - (一), 財政部。
18. 許智豪(2003), 「在動態資料庫中作線上挖掘關聯式法則」, 中興大學資訊科學研究所碩士論文。
19. 陸坤義 (2003), 「應用分層隨機抽樣和動差保留法採掘重要關聯規則之方法」, 東海大學資訊工程與科學研究所碩士論文。
20. 張文湘(2001), 「稅務行政機關服務品質及顧客滿意度之研究 - 以苗栗稅捐處為例」, 靜宜大學企業管理學系碩士論文。
21. 彰化縣稅捐稽徵處(1998), 彰化縣房屋標準單價表暨相關法令, 彰化縣稅捐稽徵處。
22. 彰化縣稅捐稽徵處(2004), 「本處簡介」、「業務資訊」, 網頁 (<http://www.changtax.gov.tw>), 彰化縣稅捐稽徵處。
23. 賴慶贊(2002), 「資料倉儲之建置與資料模式適域性之研究 - 以營業稅申報查核為例」, 銘傳大學資訊管理研究所碩士論文。
24. 蘇木春、張孝德(1999), 機器學習:類神經網路、模糊系統以及基因演算法則, 金華科技圖書。
25. 劉泓郁(2003), 「建構一個以模糊關聯法則為基礎之產品開發系統」, 元智大學工業工程與管理學系碩士論文。
26. Agrawal, R. and Srikant, R. (1994), "Fast Algorithm for Mining Association Rules," Proc.of the 20th VLDB Conference, pp.487-499.
27. Amir Netz Surajit Chaudhuri Usama Fayyad Jeff Bernhardt(2001), "Integrating Data Mining with SQL Databases: OLE DB for Data Mining", IEEE pp.379-387.
28. Baldwin, J.F. (1996), "Knowledge from data using fuzzy methods", Pattern Recognition letters, Vol.17, pp.593-600.
29. Dan Rasmussen, Ronald R. Yager(1997), "Summary SQL - A Fuzzy Tool For Data Mining", Intelligent Data Analysis, Vol.1, pp.49-58.
30. Ding-An Chiang, Louis R. Chow, Yi-Fan Wang(2000), "Mining time series data by a fuzzy linguistic summary system", Fuzzy Sets and System, Vol.112, pp.419-432.
31. Delesie, L. and Croes, L. (2000), "Operations research and knowledge discovery: a data mining method applied to health care management", International transactions in operational research, Res. 7, pp.159-170.
32. Feelders, A. and Daniels, H., Holsheimer, M. (2000), "Methodological and practical aspects of data mining", Information & Management Vol.37, pp.271-281.
33. Guoqing Chen, Qiang Wei(2002), "Fuzzy association rules and the extended mining algorithms", Information Sciences Vol.147, pp.201-228.
34. Gregory Vert, Ashley Morris, and Molly Stock, "Converting a Fuzzy Data Model to an Object-Oriented Design for Managing GIS Data Files", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 15, NO. 2, pp.510-511.
35. Hisao Ishibuchi, Takashi Yamamoto(2004), "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining", Fuzzy Sets and System Vol.141, pp.59-88.
36. James F. Smith III(2002), "Data Mining for Multi-agent Fuzzy Decision Tree Structure and Rules", International Society of Information Fusion, pp.72-79.
37. Jose Galindo, M. Carmen Aranda, Jose L. Caro, Antonio Guevara, Andrees Abuayo(2003), "Applying fuzzy databases and FSQL to the management of rural accommodation", Tourism Management, Vol.24, pp.457-463.
38. Maelainin, S.A. and Bensaid, A. (1998), "Fuzzy Data Mining Query Language", IEEE Second International Conference on Knowledge-Based Intelligent Electronic System, pp.335-342.
39. Nigel Robinson and Mary Shapcott(2002), "Data Mining Information Visualisation Beyond Charts and Graphs", IEEE Computer society, Proc of the Sixth International Conference on Information Visualisation (IV '02).
40. Oded Maimon, Abraham Kandel, Mark Last(2001), "Information-theoretic fuzzy approach to data reliability and data mining", Fuzzy Sets and System Vol.117, pp.183-194.
41. Stuart H. Rubin(1998), "A Fuzzy Approach Towards Inferential Data Mining", Computers ind. Engng, Vol.35, Nos 1-2, pp.267-270.
42. Sushmita Mitra, Sankar K. Pal, Pabitra Mitra(2002), "Data Mining in Soft Computing Framework: A Survey", IEEE Transaction on neural networks, Vol.13, NO.1, pp.3-14.
43. Tzung-Pei Hong, Kuei-Ying Lin, Shyue-Liang Wang(2003), "Fuzzy data mining for interesting generalized association rules", Fuzzy Sets and System, Vol.138, pp.255-269.
44. Usama Fayyad, Paul Stolorz(1997), "Data mining and KDD: Promise and challenges", Future Generation Computer Systems Vol.13, pp.99-115.
45. Witold Pedrycz(1998), "Fuzzy set technology in Knowledge discovery", Fuzzy Sets and System, Vol.98, pp.279-290.
46. Witold Pedrycz, and Marek(2003), "Evolutionary Fuzzy Modeling", IEEE transactions on Fuzzy System, Vol. 11, NO.5, pp.159-170.
47. Yi-Chung Hu, Ruey-Shun Chen, Gwo-Hshiung Tzeng(2003), "Finding fuzzy classification rules using data mining techniques", Pattern Recognition Letters, Vol.24, pp.509-519.
48. Yi-Chung Hu, Gwo-Hshiung Tzeng(2003), "Elicitation of classification rules by fuzzy data mining", Engineering Applications of Artificial Intelligence, Vol.16, pp.709-716.
49. Zadeh, L. A.(1965), "Fuzzy sets", Information and Control, Vol. 8, pp. 338-353.
50. Zadeh, L. A.(1975), "The concept of a linguistic variable and its application to approximate reasoning I, II, III", Information Science, 8, 199-251, 301-357; Vol. 9, pp. 43-80.