

# 發展一個最佳化之模糊關聯規則演算法

研究生：陳依伶

指導教授：蔡介元 博士

元智大學 工業工程與管理研究所

## 摘要

在資料探勘(Data Mining)技術中，關聯規則的應用是最為廣泛且最常被討論的，並被視為是適合應用於購物籃分析(Market Basket Analysis)，以展現顧客購買行為規律性的一種方法。然而，受限於資料型態的問題，過去研究所應用的資料型態大多為類別項目形式，使關聯規則探勘方法無法應用在有數值型態的資料中，為了解決上述之問題，本研究發展一個適合應用於數值型態資料的「最佳化之模糊關聯規則演算法」。

本研究所提出之最佳化模糊關聯規則演算法，是以基因染色體的觀念將資料的特徵屬性個數與期望產生的語意項目個數進行編碼，並且利用基因演算法尋找最佳解的長處，客觀且自動地產生各特徵屬性的模糊歸屬函數參數，找到適合於資料集合的模糊集合，改善以往模糊集合歸屬函數參數必須由人為設定，較不具客觀性的缺點；數值型態資料運用基因演化所產生的模糊集合進行模糊化後，可以衡量每筆數值型態資料對應於每個模糊語意項目區間的歸屬程度，並以FTDA(Fuzzy Transaction Data-mining Algorithm)演算法產生適當地模糊關聯規則，改善過去無法以數值型態資料進行關聯規則探勘的限制；最後，以測試資料驗證所挖掘到的模糊關聯規則是否具有預測準確性，並將結果做為基因演算法中染色體疊代的根據，以挖掘出適於使用者需求且預測準確性佳的模糊關聯規則。

經過一系列的實驗分析後，我們發現模糊關聯規則演算法能夠順利的從數值型態資料中挖掘到適當地模糊關聯規則，所獲得的預測準確性也能有不錯的表現。

關鍵詞：資料探勘、模糊關聯規則、基因演算法、FTDA 演算法