

## 摘要

近年來科技不斷的推陳出新，資訊科技和網際網路也迅速的發展，連帶著各式各樣的資料也快速增加，而在資料庫中隱藏著對於決策者而言有用的資訊，所以為了從這些資料中取得有用的資訊以做為決策的參考，發展出了資料探勘的技術。其中最被廣泛應用的為關聯規則。

關聯規則演算法有很多，而 Apriori 演算法是最早被提出也最具代表性的。關聯規則是以項目有沒有出現作為評斷，1 就代表有，0 就代表沒有，計算每個項目在資料出現的次數，但是在生活上很多資料型態都是數值型的，像是血壓、身高、年齡，為了讓其內容能更加精準地被表達，本研究結合模糊理論，以隸屬值做為該區間的值。

根據衛福部健康署統計糖尿病是十大死因之一，每年有將近萬人因為糖尿病而死亡，全國大約有兩百多萬名的糖尿病患者，每年都還持續增加中，糖尿病會引起其他併發症，醫療負擔不容小覷。

本研究將以 UCI 資料庫印第安皮馬人糖尿病的資料做探討，先以 FCM 和三角型隸屬函數模糊化得隸屬值，再以模糊 Apriori、Apriori 演算法及決策樹探勘規則，最後評估模型並做比較。

關鍵字：模糊理論、FCM、Apriori 演算法、決策樹