

國立高雄大學統計學研究所

113 學年度書報討論題目暨摘要登記表

結合決策樹與網絡結構的資料分群方法

張慈琪

摘要

在使用網絡進行資料分群的方法上，點與點之間的關係通常以相似度為基礎來構建，而相似度通常依賴多維距離公式進行計算。本文嘗試考慮每個點自身的特性以及與其他點之間的關聯性，針對資料分群提出了一種基於決策樹的新方法，用於建立訓練集之間的網絡關係。首先，利用決策樹模型幫助訓練集各點選取關鍵特徵及其門檻，接著建立一個代表訓練集關係的網絡，並進一步計算每個節點的重要性，以進行分群任務。完成訓練集的網絡建構後，將從訓練集獲得的特徵與門檻應用於測試集，評估此方法在測試集數據上的分群與分類表現。我們以帕金森氏症病患資料作為研究對象，初步結果顯示，此方法能有效提取數據特徵並完成訓練集的分群，但測試集的表現仍有進一步改善的空間。

關鍵詞：分群、帕金森氏症、網絡

指導教授簽名：

