

國立高雄大學統計學研究所

114 學年度書報討論題目暨摘要登記表

基於資料幾何特性之自適應混合決策樹模型研究

黃天宇

摘要

這項研究主要針對醫療診斷數據（特別是帕金森氏症）中常見的高維度、高雜訊及類別邊界模糊等特性進行探討。由於傳統決策樹雖然具備高可解釋性，但受限於其「正交切割」的特性，在處理具有斜線線性關係、非線性幾何結構（如圓形）或呈現「雲霧狀重疊」的資料時效率極差；而 SVM 或單純貝氏等模型雖然分類能力強，卻缺乏對資料進行分區處理的彈性。鑑於此，本研究提出了一種「基於資料幾何特性之自適應混合決策樹模型」，採用兩階段架構來解決上述單一模型的侷限。第一階段先利用決策樹進行全局篩選，處理能以單個特徵分割的易分類樣本；第二階段則針對低純度的「難分類葉節點」，依據資料分佈特性，改用邏輯斯迴歸、SVM 或單純貝氏進行二次建模。實驗結果經由三種幾何模擬數據及帕金森氏症臨床資料驗證。

關鍵詞：自適應混合模型、資料幾何特性、兩階段架構、類別重疊

指導教授簽名：

