

# 國立高雄大學統計學研究所

## 110 學年度書報討論題目暨摘要登記表

### 基於實驗設計準則下的熵控制拉普拉斯正規化的主動學習

#### 分類法

許傑雅

#### 摘要

模式識別是現今對於影像及圖片研究的一個重要議題，若要進行模型訓練，已標籤樣本是很重要的訊息，然則獲取大量特性或標籤需要耗費成本以藉由人工或其他方式對資料進行標記，而在大多數實例中，擁有標籤資訊的資料為少數，難以取得充足的圖像訊息，因此，透過傳統之最小平方類型的分類方法無法建構合適之分類模型。故而文獻上根據給定的流形結構之組合提出一基於熵控制的分類方法，由最大化熵懲罰調整自動化學習線性組合的對應係數。其在訓練過程中同時導入標記和未標記數據之影響性，並根據數據自動化學習流形超參數。然而，取決被標記的樣本是進一步需要探討的問題。因此，在此研究裡，我們提出一種結合實驗設計的最適化準則和文獻提出的熵控制之拉普拉斯正規化方法的主動學習法，亦即透過實驗設計準則進行標籤樣本之選取後建構出分類模型，並將方法應用於實際圖像數據集，以分析基於不同設計準則下二元分類的熵控制拉普拉斯正規化之主動學習法在圖像檢索上的分類表現。

關鍵詞：A-最適設計、D-最適設計、E-最適設計、拉普拉斯正規化、流形學習、實驗設計準則

指導教授簽名：許湘伶