資科系 53323 王人正

機器學習HW3\_Clustering

程式碼語言:

目標項:loan或y

內容:

1. 讀取金融市場行銷數據Bank.csv檔案，共有17個特徵與45,211筆樣本。
2. 取出age、balance、duration、campaign、previous五個特徵、進行標準化。
3. 以k-means進行分群
4. 透過散佈圖、轉折圖、輪廓係數、調整蘭德指數決定集群數量。
5. 以上題(a)決定的集群數量進行分群，統計每個集群的大小，並以表格、長條圖與熱度圖進行分群結果的分析說明。
6. 以ward linkage策略的凝聚分群法進行分群，統計每個集群的大小，繪製散佈圖、樹狀圖、熱度圖，並計算輪廓係數、調整蘭德指數。
7. 以DBSCAN進行分群，統計每個集群的大小，繪製散佈圖，並計算輪廓係數、調整蘭德指數。
8. 以鄰近傳播進行分群，統計每個集群的大小，繪製散佈圖，並計算輪廓係數、調整蘭德指數。
9. 以表格彙整4種分群方法的結果，並進行綜合分析。