

摘要

隨著機器學習的發展，許多先進的降維技術已逐漸被提出，而資料降維是一個減少資料分析處理難度的核心方法，此法可以進行資料視覺化和觀察資料型態與分布，但又不影響資料原本的特性，透過了解與發掘各變量的特徵，進行特徵選擇與特徵攫取，針對欲進行資料分析的目標更透徹且更有效率執行，本專案利用台灣主要縣市高鐵車站的距離矩陣與飲料資料集進行資料降維的實作；台灣主要縣市高鐵車站的距離矩陣應用 MDS 降維，飲料資料集分別應用到詞嵌入轉換後的相似度矩陣降維與 1-of-k 降維。

關鍵字：資料降維、距離矩陣、詞嵌入、相似度矩陣。