

建構一個PSO-SA的最佳化分類器於不平衡序列資料的分類

研究生：李文仁

指導教授：蔡介元教授

元智大學 工業工程與管理研究所

摘要

資料探勘中非常有名的技術「序列模式探勘」與「序列分類」的技術可以解決許多問題，例如顧客行為變化。然而，傳統的方法對於正確分類出少數樣本的能力不足。實際上，類別資料不平衡的問題經常會發生在日常生活中，例如詐騙行為的偵測、醫療診斷、垃圾信件偵測、產品監控與檢測等，因此本研究將針對不平衡序列資料發展出一個有效的分類方法。在本研究中，根據每個序列所屬的類別，AprioriAll演算法將被用來找出該類別的序列樣式，接著使用pairwise coupling方法將原始的多重類別序列資料拆成組合成許多組二類別的資料，針對每一組二類別的資料，我們會使用所提出的FMCIS方法來建構分類器並命名為FMCIS。每一個分類器會先對一條序列產生兩個相似度的值，接著再運用這兩個值去建構模糊偏好關係所需要的單元。然後使用模糊偏好關係將各個分類器所產生出來的單元值進行整合並加以計算，根據本研究所設定的終止條件將可以產生最後的分類結果。為了增加分類的準確性，一種混和的PSO-SA演算法將會被提出來調整FMCIS裡面序列樣式的權重以及模糊偏好關係裡面類別的權重。結果顯示出本研究提出的分類模型可以有效的解決序列資料不平行的分類問題，但是在模糊偏好關係這個部分，類別的權重並沒有辦法能夠有效地提升整體的分類準確率。

關鍵詞：序列樣式探勘 類別不平衡問題 序列分類方法 粒子群最佳化