

(様式7)

学位論文審査結果の要旨

氏名	加藤 聡
審査委員	委員長 伊藤 良生 印 委員 李 仕剛 印 委員 近藤 克哉 印 委員 _____ 印 委員 _____ 印
論文題目	クラスタリング問題への適用に向けた自己組織化マップの学習法改善に関する研究
審査結果の要旨	<p>審査結果の要旨</p> <p>大量のデータをグループ化し、データに内在する規則性や構造を見出す手法であるクラスタリングは、重要なデータ解析法の一つである。自己組織化マップ (SOM) は、データの分布を反映したコードベクトルが学習によって得られることから、クラスタリング問題への適用が注目されている。このとき、SOMの基本学習アルゴリズム (BSOM) では、データの密集する領域外に一部のコードベクトルが残留し、これらのコードベクトルを持った不活性セルがクラスタリングにおいて誤分類を引き起こすことが問題となる。一方、BSOMに対して、しきい値作用を導入して不活性セルの発生を抑える THSOM は、対象データの位相を保存した学習が困難なため、クラスタリングには適用できない。これらの問題を解決するため、BSOMの学習結果を THSOMの学習の初期状態とする2段階 SOMを提案した。2段階 SOMは、その学習過程において、対象データの位相を保存するとともに、不活性セルの発生を抑制することができ、これにより SOMを用いたクラスタリングにおける誤分類を抑制することができた。また、2段階 SOMに対して、クラスタ数や個々のクラスタにおける分布の特徴などを様々に変化した人工的な評価用データセットによってクラスタリング性能を詳しく検証し、クラスタごとにデータの密度が大きく異なる場合には不活性セルの判定が正しく行われず、クラスタリングの精度が低下する場合があることを示した。これは、不活性セルの判定が、隣接セル同士のコードベクトル間距離と学習時のセルの勝利回数によって行われ、これらの値がクラスタにおけるデータの密度に大きく依存するためである。この問題を解決するため、不活性セルの判定法として、隣接セル同士のコードベクトル間距離や学習時におけるセルの勝利回数の、セル間での変化量に着目した、拡張2段階 SOMを提案した。これにより、クラスタごとにデータの密度が大きく異なるようなデータセットに対するクラスタリングの精度が改善され、SOMを用いたクラスタリングの適用範囲を拡張することができた。さらに、UCI Machine Learning Databaseの実データを用いたクラスタリング性能評価により、従来の階層的・非階層的クラスタリング手法や、BSOM、2段階 SOMと比較した場合の、拡張2段階 SOMの有効性を示した。</p> <p>これらの成果は、自己組織化マップの応用、並びにクラスタ分析法の研究に新たな知見を与えるものとして評価できる。従って、本論文は博士 (工学) を授与するに値するものと認められる。</p>