

平成 17 年 1 月 12 日

機能材料 工学専攻	学籍番号	983739	指導教員氏名	高橋由雅
申請者氏名	藤島悟志			阿部英次

論文要旨(博士)

論文題目	化学構造のTFS表現による薬物構造データマイニングに関する研究 ()
------	--

(要旨 1, 200字程度)

現在、医薬や農薬の開発現場では化学構造と種々の性質との関連を積極的に見出し、これらの情報を活用した合理的な新薬開発が進められている。また一方で、例えば大きな期待を背負って登場した新薬に対して重篤な副作用が報告されたり、農薬においてはそのヒト健康や残留性による環境影響などが社会的な問題になっている。こうした中、近年の各種データベースの開発とその集積データの急激な増大に伴い、大量のデータを系統的、自動的に解析・処理しながら、有益な情報の抽出や新たな知識の発見・獲得を目的としたデータマイニングの技法が注目を集めているが、化学構造情報を基礎としたデータマイニングの研究はまだ緒についたばかりであり、その要素技術としての効果的な構造特徴解析の技法の開発やそのシステム化が喫緊の重要な課題となっている。

本研究では、薬物構造データマイニング手法の確立を目的に、先に高橋らが提案した、事前の部分構造知識を必要としない構造特徴のプロファイリング手法である Topological Fragment Spectra (TFS) 法に注目し、機械学習による薬物活性クラスの分類/予測における有用性を検証するとともに、構造類似性評価を基礎としたデータマイニングにもとづく知識発見への応用の可能性を検討した。はじめに、4種の異なる受容体に作用するドーパミン(Dn)アンタゴニストの分類/予測実験を通じてTFSを基礎とした人工ニューラルネットワーク (TFS/ANN) による活性クラス分類について検討を行った。その結果、このクラス分類においては比較的良好な結果を得たものの、大量のノイズデータ存在下での適用に際しては、学習結果がサンプル数の多いノイズに大きく影響を受けることを明らかにした。このことから、より良好な分類/予測モデルの獲得を目的として、近年、強力な分類学習モデルの一つとして注目を集めているサポートベクターマシン (SVM) の応用を試みた。TFSを入力シグナルとしたSVM (TFS/SVM) 法は、学習時における各クラスのサンプル数に大きな影響を受けることなく、何れの場合もTFS/ANN法を遥かに上回る良好な予測安定性が得られることを示した。一方、TFS表現を基礎とした構造データマイニングにおけるデータ解析と知識発見に際してのTFSピークの化学的な意味解釈を目的に、生成フラグメントやその由来構造の参照機能などを備えた、TFSピーク同定システムの開発を行った。これにより、QSARモデリングにおけるTFS記述子の意味を具体的な構造特徴と関連づけることが可能となった。さらに、化学構造間の差異の記述を目的に、新たにTFSの差スペクトルの概念を導入し、TFSピーク同定と合わせて構造特徴解析を試みた。その結果、異なる活性クラスに属する化合物間のTFS差スペクトルを解析することにより、注目する活性クラスにのみ存在する特徴的なTFSピークを容易に発見でき、かつ、そのピークを同定することにより、特徴的なフラグメントを容易に確認できることを示した。これら一連の結果より、化学構造のTFS表現は、薬物構造データマイニングにおける知識発見、および構造類似性評価にもとづく活性予測やリスク推定に対して極めて有効なアプローチを提供するものであると結論づけられる。