

平成27年1月22日

学位論文の審査要旨

学位論文申請者氏名：Relator Raissa Tillada

論文題目：Machine Learning Approaches for Biological and Physiological Data
生物学的データと生理学的データのための機械学習法

論文の概要及び判定理由

機械学習はデータ分類、データマイニングといった解析タスクのための強力なツールである。しかしながら、利用可能な情報が増えるにつれ、異なる構造のデータや異なる性質のデータが出現し、それぞれに対する高精度な解析を行うにはそれぞれに適した方法論を考案せざるを得なくなった。そのような中、生命情報学、計算化学、神経情報学のような、近年、注目を集めている分野において使われる生物学的データや生理学的データを解析するために、本論文では、それぞれに対する新しいアルゴリズムを提案し、従来法より性能が向上することを示している。具体的には次の3つの貢献が認められる。

1. 正準相関分析 (CCA) を改良して、薬剤とタンパク質の相互作用の有無を予測するアルゴリズムの性能を向上させた。薬剤タンパク質相互作用予測は薬の設計と発見において重要な役割を担う。大量に候補化合物とターゲット遺伝子がある場合、相互作用のウェット実験は時間と労力がかかりすぎる。故に、計算機的アプローチが重要視されてきた。薬剤タンパク質相互作用の予測問題の設定は2クラス分類問題になる。これまではサポートベクトルマシン (SVM) とCCAを組み合わせた方法が最もよい性能を得ていた。本研究では、CCAを拡張した、重みつきCCAを開発した。重みつきCCAによって、化合物間の類似度とターゲット蛋白質との相互作用の関係を抽出する。重みつきCCAを使ってもっとも重要な特徴を抽出して、その特徴をSVMに利用する。実験の結果、従来のSVMのみを用いた方法や、従来のCCAを用いた方法より、よい予測性能を得ることができた。

2. ブレグマンダイバージェンスを利用した学習アルゴリズムを新しく開発し、酵素の活性部位を探索する方法の性能を向上させた。酵素タンパク質における活性部位の予測はタンパク質科学において重要であるのみならず、産業応用上にも有用である。酵素反応のメカニズムは酵素活性部位の局所構造によっている。これまでは平均二乗誤差のような簡単な方法を使ってタンパク質の局所構造どうしを比較してきた。この比較能力を向上させるため、偏差にパラメータを導入する。正則化関数にブレグマンダイバージェンスを使って二乗誤差のパラメータの値を決める新しい機械学習アルゴリズムを開発した。提案法は既存の方法よりもよい探索性能を得た。

3. EEGデータのようなベクトル列で表現されるデータを解析するための新しいカーネルを考案し、従来のカーネルよりも識別性能がよいことを確認した。脳機械インターフェース (BCI)

における識別タスクは、認知科学の発展、障害者支援への応用などで注目されている。提案法は、EEG信号を使った2クラス分類のために、ベクトルの集合どうしのカーネルを使うものである。これまではグラスマン多様体によって定式化された方法が、EEG信号の識別において注目されてきた。本研究では、グラスマンカーネルに代わる新たなカーネルを開発した。このカーネルはベクトル列どうしの類似度を直接計算するもので、グラスマン射影カーネルやグラスマンビネコシカーネルのように間接的にベクトル列どうしの類似度を求める方法とは対照的である。本研究では、提案カーネルとグラスマン射影カーネルの間に理論的關係を見出し、なぜ提案法のほうがよりよい予測性能を示すのか理論的に明らかにした。

以上のように、本論文は生物学的および生理学的なデータの解析のための計算機的方法の確立に貢献する工学的価値の高い研究成果であり、また学位論文申請者は最終試験においても十分な学識を示したので、博士（工学）の学位に値するものと判定した。

審査年月日 平成27年1月20日

審査委員

| | | | | |
|----|-----------|-----|-------|---|
| 主査 | 群馬大学学術研究院 | 教授 | 関 庸一 | 印 |
| 副査 | 群馬大学学術研究院 | 教授 | 太田直哉 | 印 |
| 副査 | 群馬大学学術研究院 | 教授 | 横尾英俊 | 印 |
| 副査 | 群馬大学学術研究院 | 教授 | 小野里好邦 | 印 |
| 副査 | 群馬大学学術研究院 | 准教授 | 加藤 毅 | 印 |

関連論文

- 1 著者名 Raissa Relator, Tsuyoshi Kato, Richard Lemence
論文題目 Improved Protein-Ligand Prediction Using Kernel Weighted Canonical Correlation Analysis (カーネル重みつき正準相関分析を用いたタンパク質リガンド予測の改善)
雑誌名 IPSJ Transactions on Bioinformatics 第6巻 18頁～28頁 2013年6月
- 2 著者名 Raissa Relator, Yoshihiro Hirohashi, Eisuke Ito, Tsuyoshi Kato
論文題目 Mean Polynomial Kernel and Its Application to Vector Sequence Recognition (平均多項式カーネルとベクトル列認識への応用)
雑誌名 IEICE Transactions on Information & Systems 第E97-D巻 第7号 1855頁～1863頁 2014年7月

参考論文

- 1 著者名 Raissa Relator, Tsuyoshi Kato, Takuma Tomaru, Naoya Ohta
論文題目 Fuzzy Multiple Subspace Fitting for Anomaly Detection (異常検出のためのファジー複数部分空間フィッティング)
雑誌名 IEICE Transactions on Information & Systems 第 E97-D 巻 第 10 号
2730 頁～2738 頁 2014 年 10 月

- 2 著者名 Tetsuhiro Kakimoto, Kinya Okada, Yoshihiro Hirohashi, Raissa Relator, Mizue Kawai, Taku Iguchi, Keisuke Fujitaka, Masashi Nishio, Tsuyoshi Kato, Atsushi Fukunari, Hiroyuki Utsumi
論文題目 Automated Image Analysis of a Glomerular Injury Marker Desmin in SDT Rats Treated with Losartan (ロサルタン治療を行ったSDTラットにおける糸球体損傷マーカーデスミンの自動画像解析)
雑誌名 Journal of Endocrinology 第 222 巻 第 1 号 第 222 巻 第 1 号 43 頁～51 頁 2014 年 7 月

- 3 著者名 Yoshihiro Hirohashi, Raissa Relator, Tetsuhiro Kakimoto, Ryuta Saito, Yasushi Horai, Atsushi Fukunari, Hiroyuki Utsumi, Kinya Okada, Tsuyoshi Kato
論文題目 Automated Quantitative Image Analysis of Glomerular Desmin Immunostaining as a Sensitive Injury Marker in Spontaneously Diabetic Torii Rats (SDTラットにおける糸球体デスミン染色損傷マーカーの自動定量的画像解析)
雑誌名 Journal of Biomedical Imaging Processing 第 1 巻 第 1 号 20 頁～28 頁 2014 年 5 月

※ 掲載決定のものも記載すること