

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 20 日現在

機関番号：20103

研究種目：若手研究(B)

研究期間：2013～2016

課題番号：25730018

研究課題名(和文) 構造情報を利用したアンサンブル学習法とその大規模問題への応用

研究課題名(英文) Ensemble learning method using structure information and its application

研究代表者

竹之内 高志 (Takenouchi, Takashi)

公立はこだて未来大学・システム情報科学部・准教授

研究者番号：50403340

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：多値判別問題は判別器の学習に大きな計算コストが必要となるため、効率的に学習可能な2値判別器を多数組み合わせることで多値判別器を構成するアプローチが取られる。この枠組において、先行して提案されている各種の統合手法を特殊ケースとして含みつつ、より効率的な予測が可能な包括的な枠組みを提案した。

マルチタスク問題に対して、計算量を抑えたアンサンブル学習手法を提案した。提案手法は通常用いられるカルバックライブラー擬距離と確率モデルの組み合わせではなく、板倉斎藤擬距離と拡張擬モデルを用いた推定に基づいている。提案法の統計的性質について考察し数値実験によってその有用性を確認した。

研究成果の概要(英文)：Multiclass classification problems sometimes require huge computational cost and then a major (and efficient) approach for the problem is to integrate multiple binary classifier. In this framework, we proposed a general framework of ensemble, which includes various kinds of conventional integration-based methods as special cases.

We proposed an ensemble-based method for the Multi-task problem. The proposed method is based on a combination of the Itakura-Saito distance and an extended model, rather than the conventional combination of the Kullback-Leibler divergence and statistical models. We revealed statistical properties of the proposed method and investigated validity of the proposed method.

研究分野：統計的機械学習

キーワード：アンサンブル学習 判別分析

1. 研究開始当初の背景

データの大規模化と共に解くべき問題の複雑化が進み、「大規模な問題」を効率的に解くことが求められている。この様な目的で機械学習手法を用いた研究が盛んに行われているが、その中でアンサンブル学習が大きな役割を果たしている。例えば、多値判別問題では最適化するべきパラメータの数が少なくともクラス数に比例して増加するため、性質のよくわかっている 2 値判別問題に分解 (エンコード) し、各問題を解いた後に組み合わせる (デコード) 方が性能的・計算量的な有利であることが知られている。そこで、データを持つ構造情報を利用して大規模な問題に対して有効に働くアンサンブル学習手法を開発し、その理論的側面を明らかにすることを目的とした。

2. 研究の目的

多値判別問題を考える際にクラス数が増大した場合においても、原理的には従来手法の適用は可能であるが、問題を現実的な計算コストで高精度に解こうと思えば様々な工夫が必要となる。本研究ではデータが保持する構造情報(順序構造, クラスタ構造)に着目し、これらを利用することが可能なアンサンブル学習手法の構築を目指した。

また、関連する複数のタスクを同時に学習させることで、タスク間に共通する要因を抽出し、タスクの予測精度を向上させるのがマルチタスク学習 (転移学習) と呼ばれる問題である。マルチタスク学習も一種のアンサンブル学習と見なせるが、各タスクのノイズの構造や定義されるドメインがタスク毎に異なる。そのため、単純に従来のアンサンブル学習手法を適用しただけでは精度の良い結果を得ることはできない。そこで各タスクがどの程度の情報を共有しているのかを加味しつつ、適応的にアンサンブルするための枠組みを構築することを目的とした。

3. 研究の方法

確率モデルを用いて判別器ごとの重要度をデータから適応的に決定することで高精度の判別を可能にするアンサンブル学習による多値判別手法を基として、より柔軟に情報を扱うことが可能な枠組みを考える。構造情報を取り込むためのナイーブなアイデアとしては、疎な解が得られるような正則化を課して、不必要な小問題への分割を削除する、などが考えられる。

マルチタスク学習は一種のアンサンブル学習と見なせるが、タスク毎にノイズの構造や定義されるドメインが異なることが通常と大きく異なる点である。ここで問題となるのは、「各タスクの情報は本当に統合しても良いのか?」ということである。タスクが類似した性質を持っている場合は統合することで精度の向上が期待できるが、性質の異なるタスクの情報を統合しても害にしかない。したがって、各タスクがどの程度の情報を共有しているのかを加味しつつ、適応的にアンサンブルするための枠組みを構築する。

4. 研究成果

4.1 構造情報を利用したアンサンブル学習手法

入力ベクトル x から、クラスラベル $y \in \{1, \dots, G\}$ を予測するための多値判別手法を考える。多値判別問題の分解は符号表と呼ばれる $J \times G$ 行列 W によってなされ、 G クラスの多値判別問題が J 個の 2 値判別問題に分解される。 W の各要素は $+1, 0, -1$ のいずれかであり、 (j, k) 要素が ± 1 の時は、 j 番目の 2 値判別問題においてクラスラベル k を ± 1 とみなし、 0 の時は j 番目の 2 値判別問題においてクラス k を無視する。 j 番目の 2 値判別問題で学習した 2 値判別器を $f_j(x)$ とする。提案手法は以下の 3 つのステップからなる。

1. 2 値判別問題から学習した $f_j(x)$ から非負の関数 $q_j(y|x)$ ($j=1, \dots, J$) を構成する。
2. 構成した関数 $q_j(y|x)$ を以下のどちらか

の方法で統合する。

$$p_u(y|x) = \operatorname{argmin}_p \sum_{j=1}^J w_j D_U(p, q_j)$$

$$p_m(y|x) = \operatorname{argmin}_p \sum_{j=1}^J w_j D_U(q_j, p)$$

ただし w_j は重み, D_U はブレグマン擬距離である。

3. $\operatorname{argmax}_y p_u(y|x)$ または $\operatorname{argmax}_y p_m(y|x)$ によりラベルを予測する。

ここで関数 $q_j(y|x)$ は必ずしも確率関数である必要はない。提案手法は

1. 非負関数 $q_j(y|x)$ としてどのような関数を用いるか?
2. p_u, p_m のどちらの統合法を採用するか?

の2点の自由度があり, 多くの従来法は, 提案手法において関数 $q_j(y|x)$ を特殊な関数系に固定したものとして記述することができる。ブレグマン擬距離や重みを変更することで, 構造情報の表現や判別機に要請される性質(精度, 頑健性等)を実現することが可能となった。

また, 理論的な解析により, 最終的に多値判別に用いるために構成する条件付き確率の構成精度と, 要素として用いる2値判別器の精度の関係について明らかにした。最終的な精度は2値判別器間の相関関係により評価することが可能であり, 相関度の低い2値判別器を用いれば用いるほど多値の条件付き確率の構成精度が向上することがわかった。

4.2 マルチタスク学習

2値判別問題に対して用いられるアダプタを板倉斎藤ダイバージェンスを用いた擬似指数モデルの同定問題として特徴付け, この結果をマルチタスク学習問題に援用することで, タスク間の共有情報を利用することで性能を向上させることが可能な, 高速かつ簡易なアルゴリズムを開発した。マルチタスク学習では, 関係のないタスク

の情報を利用してしまうと精度が低下してしまう。そこで擬似指数モデルと板倉斎藤ダイバージェンスを用いてタスク間の距離を定義することで, 従来のアダプタと同様の枠組みで情報の共有の程度を制御することが可能な手法を提案した。具体的には, p_k, q_k を k 番目のタスクにおけるデータの経験分布, 擬似指数モデル, IS を板倉斎藤ダイバージェンスとして, 以下のリスク関数を最小化することで, k 番目のタスクに対する予測器を得ることができる。

$$\operatorname{argmin}_{q_k} IS(p_k, q_k) + \sum_{j \neq k} \lambda_j IS(q_k, q_j)$$

ここで λ_j は各タスクの関係度を表現するハイパーパラメータである。 λ_j を適宜調整することで, 情報を共有しているタスクのデータをより効率的に利用することが可能となり, タスク固有のノイズに対してロバストな判別器を構成することが可能となった。

5. 主な発表論文等

[雑誌論文] (計 2 件)

1. T. Takenouchi, A Novel Parameter estimation method for Boltzmann machine. *Neural computation*, 27(11), pp. 2423-2446, 2015. 査読有り.
2. T. Takenouchi, Osamu Komori and Shinto Eguchi. Binary Classification with a Pseudo Exponential Model and Its Application for Multi-Task Learning. *Entropy*, 17, pp. 5673-5694, 2015. 査読有り.

[学会発表] (計 15 件)

1. 竹之内高志. 変形ブレグマン擬距離とその周辺. 統計的モデリングと計算アルゴリズムの数理と展開, 2017. 2017年2月18日 名古屋大学 (愛知県名古屋市)

2. 竹之内高志, 変形ブレグマン擬距離とその応用. 第19回情報論的学習理論ワークショップ, 2016. 2016年11月17日 京都大学(京都府京都市)
3. 金森 敬文, 竹之内高志. グラフ上の局所ブレグマンダイバージェンスによる統計的推論. 第19回情報論的学習理論ワークショップ, 2016. 2016年11月17日 京都大学(京都府京都市)
4. 竹之内高志, 変形ブレグマンダイバージェンスを用いたパラメータ推定. 2016年度 統計関連学会連合大会. 2016年9月6日 金沢大学角間キャンパス(石川県金沢市)
5. 金森 敬文, 竹之内高志. 離散空間上のグラフ構造に基づく局所ブレグマンダイバージェンス. 2016年度 統計関連学会連合大会. 2016年9月5日 金沢大学角間キャンパス(石川県金沢市)
6. T. Takenouchi. Parameter Estimation with Deformed Bregman Divergence. Information Geometry and its Applications IV, 2016. 2016年6月13日 Liblice, Czech Republic
7. T. Takenouchi. Parameter estimation with Unnormalized model and Homogeneous divergence. Probabilistic Graphical Model Workshop: Sparsity, Structure and High-dimensionality, 2016. 2016年3月25日 統計数理学研究所(東京都立川市)
8. T. Takenouchi and T. Kanamori. Empirical Localization of Homogeneous Divergences on Discrete Sample Spaces Neural Information Processing Systems, 2015 (Spotlight). 2015年12月9日 Montreal, Canada
9. 竹之内高志, 金森敬文. 同次ダイバージェンスとその応用. 第18回情報論的学習理論ワークショップ (IBIS2015, ベストプレゼンテーション賞). 2015年11月27日 つくば国際会議場(茨城県つくば市)
10. 竹之内高志, 金森敬文. 同次ダイバージェンスを用いた離散空間上の確率モデルのパラメータ推定. 2015年度 統計関連学会連合大会連合大会. 2015年9月7日 岡山大学津島キャンパス(岡山県岡山市)
11. K. Machida and T. Takenouchi. Non-negative Matrix Factorization based on Gamma-Divergence International Joint Conference on Neural Networks (IJCNN), 2015. 2015年7月14日 Killarney, Ireland
12. 竹之内高志. α ダイバージェンスを用いたボルツマンマシンのパラメータ推定. 第17回情報論的学習理論ワークショップ, 2014. 2014年11月17日 名古屋大学(愛知県名古屋市)
13. T. Takenouchi, O. Komori, and S. Eguchi. A novel boosting algorithm for multi-task learning based on the Itakuda-Saito divergence. In BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS IN SCIENCE AND ENGINEERING (MAXENT 2014), 1641, pp. 230-237, 2014. 2014年9月25日 Amboise, France
14. 町田浩平, 竹之内高志. γ ダイバージェンスを用いた非負値行列因子分解. 情報処理北海道シンポジウム

2014. 2014年6月7日 はこだて未来大学(北海道函館市)

15. 竹之内高志. 正規化項の計算を必要としないボルツマンマシンのパラメーター推定法. 第16回情報論的学習理論ワークショップ, 2013. 2013年11月12日東京工業大学(東京都目黒区)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

○取得状況 (計 0 件)

名称:
発明者:
権利者:

種類:
番号:
取得年月日:
国内外の別:

[その他]
ホームページ等
<http://www.fun.ac.jp/~ttakashi/>

6. 研究組織

(1) 研究代表者

竹之内 高志(TAKENOUCHI, Takashi)

公立はこだて未来大学・システム情報科学部・
准教授

研究者番号: 50403340

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号:

(4) 研究協力者

()