

論文 / 著書情報
Article / Book Information

題目(和文)	クラス事前確率変化下における統計的学習 - クラス事前確率推定及びその高次元識別問題への応用 -
Title(English)	Statistical Learning under Class-prior Change: Class-prior Estimation and Its Application for High-dimensional Classification
著者(和文)	川久保秀子
Author(English)	Hideko Kawakubo
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10552号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:篠田 浩一,杉山 将,村田 剛志,下坂 正倫,鈴木 大慈,徳永 健伸
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10552号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第		号	学位申請者氏名		川久保 秀子	
論文審査 審査員		氏 名	職 名		氏 名	職 名	
	主査	篠田 浩一	教授	審査員	鈴木 大慈	准教授	
	審査員	徳永 健伸	教授		杉山 将	連携教授	
		村田 剛志	准教授				
		下坂 正倫	准教授				

論文審査の要旨（2000 字程度）

本論文は「Statistical Learning under Class-prior Change: Class-prior Estimation and Its Application for High-dimensional Classification」と題し、英文5章から成っている。

第1章「Introduction」では、機械学習の研究分野を概観するとともに、本博士論文で取り組む研究課題、および、本博士論文の全体構成を示している。機械学習の分野では、これまで主に教師付き学習と教師なし学習の研究が行われてきており、最近はその中間に位置づけられる半教師付き学習も盛んに研究されている。このような機械学習の枠組みにおいて、データを生成する確率分布が変化する状況にいかにもく適応するか、また、高次元のデータからいかに精度良く学習を行うかが、近年は特に重要な研究課題となっている。本章では、標準的な半教師付き分類問題を数学的に定式化した後、訓練時とテスト時でデータの生成確率分布が変化する状況での分類問題を考え、共変量シフト、および、クラス事前確率変化とよばれる問題設定における適応学習法の概要を紹介している。次に、高次元データの次元削減問題を考え、入力空間のうち出力の予測に十分である低次元の部分空間を見つける十分次元削減とよばれる枠組みを紹介している。最後に、本博士論文の目的は、高次元データを対象とした、クラス事前確率が変化する状況下での半教師付き分類問題に対する新しい機械学習アルゴリズムを提案することであると述べている。

第2章「Review of Existing Methods and Our Strategies」では、分類問題に対する従来の様々な手法を紹介している。まず、分布適合とよばれる、クラス事前確率変化に対する適応手法が必要となるクラス事前確率推定の枠組みを紹介している。そして、この枠組のもと、カルバック・ライブラー距離や最大平均差異などの確率分布間の距離尺度に基づくクラス事前確率推定手法を紹介し、推定の際にパラメータの調整に時間がかかるという問題、あるいは、パラメータの調整が客観的に行えないという問題があることを指摘している。次に、エネルギー距離とよばれる確率分布間の距離尺度を紹介している。エネルギー距離の推定の際には調整すべきパラメータがないため、その効率性を活かした手法を第3章で提案する。次に、十分次元削減の従来法である、逆回帰や回帰関数のヘシアンを用いた手法、条件付き確率推定に基づく手法、相互情報量に基づく手法、カーネルに基づく手法、最小二乗勾配推定に基づく手法を紹介している。そして、クラス事前確率変化のもとでは、これらの十分次元削減手法は性能が大幅に低下してしまう事を示している。これら従来の十分次元削減法の中では、最小二乗勾配推定に基づく手法の性能が良いことが知られているため、第4章ではこれをクラス事前確率変化に適応できるように拡張する。

第3章「Semi-supervised Class-prior Estimation via Energy Distance」では、エネルギー距離を用いたクラス事前確率の推定手法を提案している。提案手法は分布適合の枠組みに則っており、エネルギー距離を用いて分布の適合を行う。提案法では、クラス事前確率推定のための目的関数が凸関数になり、この凸最適化問題が解析的に解けることを示している。計算機実験によって、提案法は従来法と同程度の推定精度を維持しつつ、計算時間が大幅に短縮できることを示している。

第4章「Semi-supervised Sufficient Dimension Reduction」では、クラス事前確率変化のもとで有用な十分次元削減手法を提案している。提案法は、最小二乗勾配推定に基づく手法に、第3章で提案したクラス事前確率推定法を組み合わせたものとなっている。計算機実験により、計算時間をほとんど増加させることなく、クラス事前確率変化のもとでの次元削減の性能とその後の分類の性能を大幅に改善できることを示している。

第5章「Conclusions and Future Work」では、本論文の成果を総括し、今後の課題を述べている。
以上を要するに本論文は、機械学習分野におけるクラス事前確率変化に対する適応学習技術を発展させるものであり、工学上、及び、工業上貢献するところが大きい。よって我々は、本論文が博士（工学）の学位論文として十分価値あるものと認める。