【8】

氏名	林（はやし） 賢一（けんいち）
博士の専攻分野の名称	博　士（工学）
学位記番号	第　24909　号
学位授与年月日	平成23年9月20日
学位授与の要件	学位規則第4条第1項該当 基礎工学研究科システム創成専攻
学位論文名	Boosting Methods for Asymmetric Mislabeled Data （非対称ミスラベルデータに対するブースティング法）
論文審査委員	（主査） 教　授　狩野　裕 （副査） 教　授　白旗　慎吾　　教　授　内田　雅之

論文内容の要旨

本論文では，非対称ミスラベルデータに対するブースティング法を提案し，その統計学的性質を調べた．ブースティングは，判別問題に対して機械学習の分野で提案された方法論である．ブースティングは判別精度の高い判別規則を構成することができる一方で，データに混入したノイズに敏感であり，そのため汎化誤差が増大するという問題がある．非対称ミスラベルモデルはノイズの混入を表現するモデルの一つである．その特徴は，観測個体のラベルが誤って観測される（ミスラベル）確率があり，その確率が真のラベルに依存して変化するという点である．非対称ミスラベルモデルは外れ値混入モデルも表現することができる．

1章では，本研究の背景としてブースティングの統計学的研究とミスラベルデータのモデリングについて概観し，問題点と本研究の目的を明確にした．

2章では，特徴量空間においてミスラベル確率が一定である非対称ミスラベルモデルを考え，Asymmetric LogitBoostを提案した．また，その近似としてAsymmetric AdaBoostを導出し，これが真のラベルに対するBayes判別境界を与えうることを示した．さらに，Asymmetric AdaBoostが形式の上でcost-sensitive学習と同等であることを示した．

3章では，より一般的なミスラベルメカニズムとして，ミスラベルの確率が共変量に依存して変化するモデルを考え，Asymmetric Eta-Boostを提案した．これはEta-Boost，AdaBoostの一般化とみることができる．提案アルゴリズムは真のラベルに対するBayes判別境界を構成することができ，これが一般化されたKullback-Leibler情報量の最小化と同等であることを示した．

4章では，罰則付きリスクを最適化するブースティングを考え，そのリスク一致性を証明した．このブースティングは，過学習を防ぐための方法として有用である．過学習とは，見かけの判別精度と汎化誤差の乖離が大きくなる現象である．ブースティングにおいてデータに外れ値が混入している場合，その影響により過学習を起こしやすい．

論文審査の結果の要旨

申請者は，提出論文において，非対称ミスラベルデータに対するブースティング法を提案しその統計学的性質を調べている．ミスラベルとは二値の従属変数の値が入れ替わる不正確なデータ（エラー，ノイズ）を指し，それは，伝統的な統計学における外れ値のごとく，統計解析を不正確にする．ブースティングは，データの非線型構造を探索的に同定することができる現代的かつ有効な手法であるが，機械学習の分野で発展してきたという歴史的な経緯もあり，ランダムエラーに弱いという欠点がある．非対称ミスラベルモデルはノイズの混入を表現するモデルの一つである．その特徴は，観測個体のラベルが誤って観測される（ミスラベル）確率が真のラベルに依存して変化する点である．

申請者は，論文第２章において，特徴量空間におけるミスラベル確率が一定である非対称ミスラベルモデルを考えAsymmetric LogitBoostを提案した．その近似としてAsymmetric AdaBoostを導出し，これが真のラベルに対するBayes判別境界を与え得ることを示した．さらに，Asymmetric AdaBoostが形式的にはcost-sensitive学習と同等であることを示した．第３章では，より一般的なミスラベルメカニズムとして，ミスラベルの確率が共変量に依存するモデルを考えAsymmetric Eta-Boostを提案した．これはEta-Boost，AdaBoostの一般化である．提案アルゴリズムは真のラベルに対するBayes判別境界を構成することができ，そして，これが一般化されたKullback-Leibler 情報量の最小化と同等であることを示した． 第４章では，罰則付きリスクを最適化するブースティングのリスク一致性を証明した．このブースティングは，見かけの判別精度と汎化誤差の乖離が大きくなるいわゆる過学習を防ぐための方法として有用である．

以上の研究成果は，統計科学において独創性に富み，応用において有益な統計手法を提供すると評価される．それゆえ，博士（工学）の学位論文として価値のあるものと認める．