

氏名	福地 一斗		
学位の種類	博士(工学)		
学位記番号	博甲第8519号		
学位授与年月日	平成30年3月23日		
学位授与の要件	学位規則第4条第1項該当		
審査研究科	システム情報工学研究科		
学位論文題目	Theoretical Analyses of Learning under Fairness (公平性下における学習の理論的解析)		
主査	筑波大学 教授	博士(工学)	佐久間 淳
副査	筑波大学 教授	理学博士	北川 博之
副査	筑波大学 教授	博士(工学)	福井 和広
副査	筑波大学 准教授	博士(工学)	日野 英逸
副査	産業技術総合研究所 主任研究員	博士(情報学)	神嶌 敏弘

論文の要旨

審査対象論文は、公平性下における学習の、公平性を実現する学習手法およびその理論的保証に関する検討を行っている。

第1章は導入である。

第2章では、経験損失最小化を導入するとともに、公平性下における学習に関する既存研究を紹介している。

第3章では、公平性を考慮すべきセンシティブ属性は直接観測できないが、モデルベースで公平性を考慮すべきセンシティブ属性が与えられる場合における、公平性下における学習を実現するアルゴリズムと、実験結果によるその有効性が述べられている。

第4章では、経験損失最小化を基にした中立化の新しい枠組みである Neutralized ERM (NERM) を提案している。既存研究は、どの手法も与えられた事例について経験的に計算される予測の中立性を実現するが、未知事例に対する中立性の保証は無かった。NERM では、未知事例に対する中立性を保証する、汎化中立性の確率的バウンドに関する理論的解析を行い、これを理論的に保証している。

第5章では、中立性の評価指標となる加法分解可能なスカラー汎関数について、その minimax 最適な推定量に関する理論解析を与えている。その minimax レートが、対象となるスカラー汎関数の divergence speed とよばれる量によって特徴づけられることを示し、それぞれについて具体的な収束レートを導出している。

審査の要旨

【批評】

審査対象論文は、公平性下における学習の、公平性を実現する学習手法について、理論的保証の検討および実験によるその有用性の評価を行っている。論文は第3章、第4章、第5章が公平性下における学習に関連する、それぞれひとまとまりの成果である。

第3章では、公平性を考慮すべきセンシティブ属性は直接観測できないが、モデルベースで公平性を考慮すべきセンシティブ属性が与えられる場合において、公平性下における学習を実現するアルゴリズムを提案し、実験結果によるその有効性が述べられた。現実問題としては、公平性を考慮すべきセンシティブ属性が直接観測できない(例えば履歴書に性別、人種、宗教を記載しない)が、それによって行われた判断(例えば雇用判断)が結果的にセンシティブ属性に依存することがありえる。この研究は、これまで取り扱われていなかった、観測することができないセンシティブ属性に対する差別を解消する機械学習アルゴリズムを与えたという意味で、新規性の高い研究である。また、その有効性は実験的に検証されている。

第4章では、公平性下における学習において、観測事例のみならず、未観測事例についての、予測のセンシティブ属性に対する中立性の保証に関する理論解析と、それを実現するための機械学習アルゴリズムが提案された。未観測事例についての中立性の保証は、機械学習理論に NERM という新しい学習理論のフレームワークを与えると意味で貢献があり、かつ、サンプル数を多くすればより厳密な中立性を保証できることを理論的に保証しているという意味で、重要な知見を伴う成果である。さらに、提案されたサポートベクターマシンをベースとした中立化を実現する分類アルゴリズムは、既存手法よりも、予測精度および中立性という意味で優れていることが実験的に示されており、有用性が高い手法であることが示された。

第5章では、中立性の評価指標となるエントロピー様の統計量(加法分解可能なスカラー汎関数)について、その minimax 最適な推定量に関する理論解析を与えた。minimax レートが、対象となるスカラー汎関数の divergence speed とよばれる量によって特徴づけられるという結果は、これまでに知られていなかった知見であり、 f エントロピーや様々なダイバージェンスの推定において、個別的に導出されていた minimax 最適な推定量を、統一的に取り扱うことができる可能性を示唆しているという意味で、理論上重要な成果である。

【最終試験の結果】

平成30年1月31日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士(工学)の学位を受けるに十分な資格を有するものと認める。