

審査の結果の要旨

氏 名 石井 雅人

機械学習アルゴリズムは目覚ましい発展を遂げ、物体認識、音声認識など、様々な実応用において高い性能を達成してきている。機械学習の高い性能を支える要因の一つが大規模な学習データセットの活用であり、データセットの重要性が増している。しかしながら、大規模なデータセットを構築するためには、多大な労力と金銭的成本がかかり、学習データ構築のためのコストがさらなる実応用への大きな障壁になることが問題視されている。その様な背景のもと、ドメイン適応技術が提案され、既存のデータセットを効果的に利用することで新たな問題に対する学習データセットを取得するコストを大幅に削減することができる技術として注目を集めている。本博士論文では、ドメイン適応の適用範囲を広げ、実社会の多様な問題に利用可能にすることを目的として、3つの手法を提案している。

本論文は以下のように全6章からなる。

第1章は序論である。本論文の対象であるドメイン適応の背景、ドメイン適応の研究分野における本博士論文の貢献および本論文の構成について述べている。

第2章では、ドメイン適応の関連研究を紹介し、特に、ドメイン適応における研究の流れと既存研究の紹介と本博士研究の当該分野における位置付けについて述べている。

第3章から第5章では、ターゲットデータが一部得られないという、ターゲットデータが全く得られない、ソースデータが全く得られないという3つの問題設定に取り組み、それぞれに対するドメイン適応手法を提案している。

第3章では、ターゲットデータが一部得られない場合のドメイン適応手法を検討し、部分的ゼロショットドメイン適応の問題、すなわち、いくつかのクラスがターゲットデータに含まれていないという問題に取り組んでいる。本章では、ターゲットデータに含まれていないクラスの未知データに対する損失を推定するため、従来の敵対的学習に基づくアプローチに対してサンプル重みづけを併用するという発想に基づく手法を提案し、複数のベンチマークデータを用いた実験により、提案手法の有効性を示している。

第4章では、ターゲットデータが全く得られない場合のドメイン適応手法を提案している。ここでは、属性と呼ぶ特定の要因の事前分布の違いに起因してドメイン間のデータ分布がずれている状況を対象として、属性の事前分布の変化が知識として得られるとい

う仮定のもとドメイン適応を行う手法を提案し、人工データとベンチマークデータを用いた実験により、提案手法のドメイン適応の有効性を検証している。

第5章では、ソースデータが全く得られないという極端な場合におけるドメイン適応の問題に取り組んでいる。具体的には、ソースデータを用いて事前学習されたモデルがソースデータの代わりに与えられるとして、このモデルをラベルなしターゲットデータのみを用いてターゲットドメインに適応する手法を提案している。複数のベンチマークデータを用いた実験により、提案手法は適応時にソースデータへのアクセスを必要としないにもかかわらず、高い性能を達成できることを示している。

第6章では、結論と今後の課題について述べている。

以上要するに、本博士論文は、不十分なデータに基づくドメイン適応について、ドメイン適応の適用可能範囲を広げるために、1. ターゲットデータが一部得られないという問題設定、2. ターゲットデータが全く得られない問題設定、3. ソースデータが全く得られないという問題設定におけるドメイン適応手法を実現している。これらの成果は、ドメイン適応の適用範囲を広げ、ドメイン適応を実問題の多様な状況に適用可能にしたという意味において、コンピュータ科学、特に機械学習分野に対する貢献が大きいものと判断される。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。