

二国間交流事業 共同研究報告書

令和5年4月24日

独立行政法人日本学術振興会理事長 殿

[日本側代表者所属機関・部局]
東京大学・先端科学技術研究センター
[職・氏名]
教授・原田達也
[課題番号]
JPJSBP120207705

1. 事業名 相手国: インド (振興会対応機関: DST) との共同研究

2. 研究課題名

(和文) 限られたラベル付きデータからの学習: 次世代の機械学習問題の解決

(英文) Learning with Limited Labeled Data: Solving the Next Generation of Machine Learning Problems

3. 共同研究実施期間 2020年7月1日 ~ 2023年3月31日 (2年9ヶ月)【延長前】 2020年7月1日 ~ 2022年6月30日 (2年 ヶ月)

4. 相手国側代表者(所属機関名・職名・氏名【全て英文】)

Indian Institute of Technology, Hyderabad, Associate Professor, Vineeth N Balasubramanian

5. 委託費総額(返還額を除く)

本事業により執行した委託費総額		1,950,000 円
内訳	1年度目執行経費	700,000 円
	2年度目執行経費	950,000 円
	3年度目執行経費	300,000 円

6. 共同研究実施期間を通じた参加者数(代表者を含む)

日本側参加者等	6名
相手国側参加者等	1名

* 参加者リスト(様式 B1(1))に表示される合計数を転記してください(途中で不参加となった方も含め、全ての期間で参加した通算の参加者数となります)。

7. 派遣・受入実績

	派遣		受入
	相手国	第三国	
1年度目	0	0	0(0)
2年度目	0	0	0(0)
3年度目	0	0	0(0)

* 派遣・受入実績(様式 B1(3))に表示される合計数を転記してください。

派遣: 委託費を使用した日本側参加者等の相手国及び相手国以外への渡航実績(延べ人数)。

受入: 相手国側参加者等の来日実績(延べ人数)。カッコ内は委託費で滞在費等を負担した内数。

8. 研究交流の概要・成果等

(1)研究交流概要(全期間を通じた研究交流の目的・実施状況)

IIT-H のチームと東京大学のチームの研究交流により、限られたラベル付きデータからの学習手法の開発の観点から、次世代の機械学習問題の解決を目指す。研究交流期間である 2020 年度から 2022 年度は、コロナウィルス蔓延の影響を最も受けた時期にあたり、国際的な交流のみならず、国内における人的交流がほとんど行えない状況となり、予定通りに研究交流を進めることが不可能であった。しかしながら、この非常に困難な状況においても、ミニワークショップの開催や IIT-H 向けの特別講義などを行えたのは大きな成果である。

(2)学術的価値(本研究交流により得られた新たな知見や概念の展開等、学術的成果)

限られた教師データしか存在しない状況での予測モデルの性能向上のために、教師なしドメイン適応について IIT-H と東京大学のチームで議論を行った。教師ラベルが大量に得られるドメイン(ソースドメイン)で学習したモデルを、ラベル情報が存在しない異なるドメイン(ターゲットドメイン)に適用した場合の性能低下をできる限り防ぐ研究領域を教師なしドメイン適応(UDA)と呼ぶ。近年のデータプライバシーに関する新たな規制により、将来的に UDA を適用する際にソースデータを利用できる保証はない。そこで、本研究交流を通じて、ソースデータを利用しない新たな UDA 問題に取り組む意義を確認し、その新規手法構築のためのアイデアを議論した。また、ドメイン適応の理論的研究は、従来無視されてきた結合誤差を考慮することが重要であることを確認し、今後進むべき方向性を定め研究を実施した。

(3)相手国との交流(両国の研究者が協力して学術交流することによって得られた成果)

東京大学のチームは、ドメイン適合とクロスドメイン生成モデル、つまり、教師情報が大量に得られる環境から教師情報がほとんど得られない制約のあるターゲット環境への学習された知識の転送に焦点を当てている。一方、IIT-H のチームは少数ショット/ワンショット/ゼロショットの学習などの設定に取り組んでいる。IIT-H と東京大学の双方のチームとも限られたラベル付きデータからの学習問題に取り組んでおり、目的を一にするチーム同士が真剣に議論をすることで、(2)で述べたようにデータプライバシーを意識した手法の開発や、UDA で根本的に見落とされてきた結合誤差に焦点を当てた手法の開発に着手できたことは大きな成果である。

(4)社会的貢献(社会の基盤となる文化の継承と発展、社会生活の質の改善、現代的諸問題の克服と解決に資する等の社会的貢献はどのようにあったか)

今後 AI は社会の必要不可欠な基盤技術となることは間違いない。AI がその高い性能を示すためには、質の高い教師データが大量に必要であり、この教師データの作成が様々な分野への AI の浸透の大きな障壁となっている。本プロジェクトはこの AI の本質的な課題を解決し、次世代の機械学習手法構築をすることで社会貢献を行うものである。

(5)若手研究者養成への貢献(若手研究者養成への取組、成果)

ミニワークショップにおいては IIT-H と東大のチームからそれぞれ 2, 3 名の学生が発表を行い活発な議論を

おこなった。また, IIT-H の学生や若手研究者に向けて東大チーム PI の原田が特別講義を実施し, 本分野における若手研究者養成へ貢献した。

(6)将来発展可能性(本事業を実施したことにより、今後どのような発展の可能性が認められるか)

本事業で目指す技術は, 大量の教師データが必要とする現状の AI の根本的な問題点を解決することで, 幅広い学術分野における新たな知見の発見に貢献可能な技術へ発展させられる可能性がある。

(7)その他(上記(2)～(6)以外に得られた成果があれば記載してください)

例: 大学間協定の締結、他事業への展開、受賞など

特になし。