郭亚明

Phone：(+86) 19838991231 · Email：yamingguo98@outlook.com

教育背景

内蒙古大学, 应用数学, 本科 2017.09 - 2021.06

- **本科成绩**: GPA：3.6/4.0 (Top 3)
  数学分析 (99)　实变函数 (100)　运筹学 (98)　数值逼近 (99)
- **英语情况**: CET4
- **获奖情况**:

1. 2019 年 “深圳杯” 数学建模挑战赛国家三等奖 (**全国 Top 5**, 队长)
2. 2020 年美国大学生数学建模竞赛 “H” 奖 (队长)
3. 国家励志奖学金, 内蒙古大学本科生优秀学生二等奖学金

吉林大学, 数学, 硕士 (保研) 2021.09 - 2024.06

- **研究兴趣**: 联邦学习, 分布外泛化, 持续学习
- **论文情况**:

1. Cao X, **Guo Y**, et al. When an Active Learner Meets a Black-box Teacher. arXiv preprint arXiv:2206.15205, 2022.
2. **Guo Y**, Guo K, et al. Out-of-Distribution Generalization of Federated Learning via Implicit Invariant Relationships. International Conference on Machine Learning, 2023. (**CCF A**)

论文介绍

- **论文 1 (导师一作)**: 主动学习通过最大化假设更新来寻找所需的未标记数据. 这种学习方式的一个内在假设是它可以将这些更新导向最优假设. 然而, 如果这些增量更新是负面的且无序的, 其收敛可能无法很好地保证. 在本文中, 我们引入了一个机器教师, 他为主动学习者提供一个黑盒教学假设作为最优假设的有效近似. 理论上, 我们证明了在这种教学假设的指导下, 学习者可以收敛到比那些没有从教师那里获得任何指导的非教育学习者更紧的泛化误差和标签复杂度界. 我们进一步考虑了两种教学场景：教授白盒和黑盒学习者, 并首次提出了自我改进的教学方法来提高教学性能. 实验证实了我们的想法, 并展示了比基本的主动学习策略 (如 IWAL, IWAL-D 等) 更好的性能.
- **论文 2**: 对于联邦学习中存在分布偏移的非参与客户而言, 分布外泛化具有挑战性. 一个行之有效的策略是探索输入与目标变量之间对非参与客户同样有效的不变关系. 然而, 学习不变关系通常是以显式的方式从数据、表示、分布中进行的, 违反了隐私保护和有限通信的联邦原则. 在本文中, 我们提出了 FedIIR, 它从参数中隐式地学习不变关系以进行分布外泛化, 同时遵循上述原则. 具体来说, 我们利用预测分歧来量化不变关系, 然后通过客户间梯度对齐隐式地减少它. 理论上, 我们给出了 FedIIR 能够推广到的非参与客户的范围, 并证明了 FedIIR 在客户端大规模分布且通信有限场景下的收敛结果. 广泛的实验表明 FedIIR 在联邦学习场景中的分布外泛化方面明显优于相关基线.

项目经历

- 2018.05-2020.04 大学生创新训练计划, **成果以论文 ‘数值域可加性研究’ 发表在 “理论数学”**, 主持人. 主要负责可加性定理的构造和证明: 利用数值域的酉不变性和正规矩阵数值域恰为其特征值的凸组合等性质, 从 Hermite 矩阵入手, 给出了二维正规矩阵数值域可加性成立的充要条件.
- 2019.04-2019.08 “深圳杯” 数学建模挑战赛, **国家三等奖 (Top 5)**, 队长. 主要负责数学模型的建立与求解: 从多目标优化视角分析了波束设计问题, 建立了逆向的波束设计模型, 并利用离散差分进化算子对 MOEA/D 算法改进以求解模型, 最后通过大量仿真和实际应用验证了模型的有效性.
- 2023.01-2025.12 面向黑盒的机器教学收敛研究, 国家自然基金项目, 参与者. 主要负责黑盒假设剪枝的可行性研究: 引入了一个黑盒机器教师, 以主动学习作为剪枝方式, 分析了黑盒机器教学的理论优势.

掌握技能

- **数学专业基础**: 泛函分析 & 机器学习
- **编程语言基础**: Python & Matlab
- **论文写作基础**: Latex & Adobe
- **深度框架基础**: Pytorch & DomainBed

Yaming Guo

Phone : (+86) 19838991231 · Email : yamingguo98@outlook.com

EDUCATION

Inner Mongolia University, Bachelor of Applied Mathematics 2017.09 - 2021.06

- **Academic Performance**: GPA : 3.6/4.0 (Top 3)
  Mathematics Analysis (99) Real Analysis (100) Operations Research (98) Numerical Approximation (99)
- **English**: CET4
- **Achievements**:

1. National Third Prize (**Top 5**) in the 2019 "Shenzhen Cup" Mathematical Modeling Challenge (Team Captain).
2. "H" Award in the 2020 American Mathematical Contest in Modeling (Team Captain).
3. National Inspirational Scholarship, Second-class Scholarship for Outstanding Undergraduates at IMU.

Jilin University, Master of Mathematics (*Recommended for Admission)* 2021.09 - 2024.06

- **Research Interests**: Federated Learning, Out-of-Distribution Generalization, Continuous Learning
- **Papers**:

1. Cao X, **Guo Y**, et al. When an Active Learner Meets a Black-box Teacher. arXiv preprint arXiv:2206.15205, 2022.
2. **Guo Y**, Guo K, et al. Out-of-Distribution Generalization of Federated Learning via Implicit Invariant Relationships. International Conference on Machine Learning, 2023. (**CCF A**)

Paper Descriptions

- **Paper 1**(Mentor as first author): Hypothesis-pruning maximizes the hypothesis updates for active learning to find those desired unlabeled data. An inherent assumption is that this learning manner can derive those updates into the optimal hypothesis. However, its convergence may not be guaranteed well if those incremental updates are negative and disordered. In this paper, we introduce a machine teacher who provides a black-box teaching hypothesis for an active learner, where the teaching hypothesis is an effective approximation for the optimal hypothesis. Theoretically, we prove that, under the guidance of this teaching hypothesis, the learner can converge into a tighter generalization error and label complexity bound than those non-educated learners who do not receive any guidance from a teacher. We further consider two teaching scenarios: teaching a white-box and black-box learner, where self-improvement of teaching is firstly proposed to improve the teaching performance. Experiments verify this idea and show better performance than the fundamental active learning strategies, such as IWAL, IWAL-D, etc.
- **Paper 2**: Out-of-distribution generalization is challenging for non-participating clients of federated learning under distribution shifts. A proven strategy is to explore those invariant relationships between input and target variables, working equally well for non-participating clients. However, learning invariant relationships is often in an explicit manner from data, representation, and distribution, which violates the federated principles of privacy-preserving and limited communication. In this paper, we propose FedIIR, which implicitly learns invariant relationships from parameter for out-of-distribution generalization, adhering to the above principles. Specifically, we utilize the prediction disagreement to quantify invariant relationships and implicitly reduce it through inter-client gradient alignment. Theoretically, we demonstrate the range of non-participating clients to which FedIIR is expected to generalize and present the convergence results for FedIIR in the massively distributed with limited communication. Extensive experiments show that FedIIR significantly outperforms relevant baselines in terms of out-of-distribution generalization of federated learning.

Project Experiences

- 2018.05-2020.04 National College Students' Innovation and Entrepreneurship Training Program; **published a paper titled 'On the Additive Property of Numerical Ranges' in "Pure Mathematics"**; responsible for the construction and proof of additivity theorems; leader.
- 2019.04-2019.08 "Shenzhen Cup" Mathematical Modeling Challenge; **National Third Prize (Top 5)**; responsible for mathematical model establishment and solution; captain.
- 2023.01-2025.12 Research on Convergence of Black-box Machine Teaching; National Natural Science Foundation of China; responsible for feasibility research on black-box hypothesis pruning; participant.

Skills

- **Mathematical**: Functional Analysis & Machine Learning
- **Coding**: Python & Matlab
- **Writing**: LaTeX & Adobe Illustrator
- **Framework**: PyTorch & DomainBed