Review of: "Causality in Machine Learning: Innovating Model Generalization through Inference of Causal Relationships from Observational Data"

Yi Zhang[1]

1 Universität Augsburg

Potential competing interests: No potential competing interests to declare.

This research paper summarizes the methods for causality inference from observational data and proposes the generalization ability of causality inference methods in machine learning. The paper is insightful for further development of machine learning algorithms. However, "synthesizing causal discovery and machine learning" remains mostly an intention statement. Neither detailed approaches nor concrete results nor quantitative comparisons of various causality inference methods are provided, resulting in a lack of trustworthiness. The following comments could be considered to improve this paper.

1. In the section literature review, the causality inference methods are categorized into causality discovery methods, integrating causality into machine learning, and causal representation learning. But in abstract and in results, three types of methods are mentioned: constraint-based, score-based, and neural structure learning algorithms. What is the relationship between them? Besides, the methods in results (like GNN, GDS) are not totally introduced in literature review. I suggest summarizing all methods into a table, including the categories, characteristics, as well as pros and cons of the methods.
2. In the sub-section theoretical frameworks, which methods mentioned before belongs to these frameworks? And how do they relate to the proposed methods?
3. In the last paragraph in the section literature review, the Reichenbach's common cause principle and the causal Markov condition are not explained.
4. In results, the numbers in Table 1 are not explained.
5. The results are analyzed, and the conclusions are drawn directly without reference to any experiments, criteria, and conditions. It contradicts the "quantitative study" claimed in introduction. As the outcomes of causality inference methods, the directed causal graphs are not shown.