# TODO

Kenny Chiu

September 20, 2021

# 1 Conceptual summary

The paper by Mouli and Ribeiro [MR21] examines the problem of extrapolating patterns learned from training data collected from a single environment to data collected from other environments. This problem context falls under the idea of *domain adaptation* that has been explored in recent literature [Far+20]. However, a key assumption in Mouli and Ribeiro's work that distinguishes it from previous work in the literature is that the training data come from a single environment as opposed to multiple environments. Several previously proposed methods for domain adaptation—such as *Invariant Risk Minimization* [Arj+20] (IRM)—rely on training data from multiple environments and therefore would fail under this problem context. Mouli and Ribeiro take a different approach by viewing extrapolation as counterfactual reasoning in a specified structural causal model (SCM) and assuming known (linear automorphism) group structures on the non-causal mechanisms. Under this formulation, Mouli and Ribeiro introduce a learning framework for the single-environment context that is able to learn invariances that do not contradict the data. In this conceptual summary, we review the key contributions of the paper by Mouli and Ribeiro [MR21] and discuss the strengths and weaknesses of their work.

## 1.1 Key differences from previous work

Various methods for domain adaptation have been proposed in the literature, and how the work by Mouli and Ribeiro [MR21] relates to these methods are highlighted in their paper. For example, methods based on causal inference such as IRM and *Independent Causal Mechanisms* [Par+18] broadly involve learning some internal representation of the data that is invariant to environment-specific, non-causal mechanisms. The invariant representation is learned from the training data which come from multiple environments. When the data come from a single environment, the representation cannot determine which aspects of the data are environment-specific and so the representation is unlikely to extrapolate to new environments. The learning framework proposed by Mouli and Ribeiro which does work with single-environment data has an advantage over existing methods in these settings.

Another approach for domain adaptation is based on data augmentation [CDL20] where training is done with not only the original data but also proper transformations of the data. Mouli and Ribeiro [MR21] explains that data augmentation is a type of *forced invariance* where certain transformations of the data may actually introduce contradictions (e.g., images of digits 6 and 9 are not invariant to $180^o$ rotations). Their proposed learning framework aims to learn only the invariances that do not contradict the data.

## 1.2 Main contributions

TODOinclude empirical results

## 1.3 Limitations

TODOlimitation: known groups

# 2 Technical summary

# 3 Mini-proposals

## 3.1 Proposal 1: MY PROPOSAL TITLE

## 3.2 Proposal 2: MY OTHER PROPOSAL TITLE

# 4 Project report

# References


[Arj+20] M. Arjovsky et al. *Invariant Risk Minimization*. 2020. arXiv: 1907.02893 [stat.ML].

[CDL20] S. Chen, E. Dobriban, and J. Lee. "A Group-Theoretic Framework for Data Augmentation". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 21321–21333. URL: https://proceedings.neurips.cc/paper/2020/file/f4573fc71c731d5c362f0d7860945b88-Paper.pdf.

[Far+20] A. Farahani et al. *A Brief Review of Domain Adaptation*. 2020. arXiv: 2010.03978 [cs.LG].

[MR21] S. C. Mouli and B. Ribeiro. *Neural Networks for Learning Counterfactual G-Invariances from Single Environments*. 2021. arXiv: 2104.10105 [cs.LG].

[Par+18] G. Parascandolo et al. *Learning Independent Causal Mechanisms*. 2018. arXiv: 1712.00961 [cs.LG].