
CS7370 : Causal Inference

Assignment 2

21th Oct 2019

Deadline: Nov 5 2019 11:55 pm

Max Marks: 10

Instructions:

- Submit the assignment in latex. Collaborations and discussions are encouraged.
 - Be precise with your explanations. Unnecessary verbosity will be penalized. Please start early.
 - Check the Moodle discussion forums regularly for updates regarding the assignment.
 - * marked questions are hard
 - You will be evaluated based on a Viva
-

1. Write summaries for the following topics.

1. (2 marks) Fairness in data, and machine learning algorithms is critical to building safe and responsible AI systems from the ground up by design. Summarize the key concepts discussed in the [book](#) related to classification (Ch. 2) and causality (Ch. 3). Short summary of introduction (Ch. 1) and basics of causality (till Graphical discrimination analysis of Ch. 3) is give [here](#).
 2. (2 marks) Removal of bias from the algorithm is an important aspect to be considered before achieving algorithmic fairness. Concepts learnt in class related to Potential Outcome(PO) framework and Pearl's approach can be used to remove the data bias. Summarize the key concepts of [Silvia et al.](#) which tries to assume hierarchy amongst features of dataset using causal diagrams and quantify data bias.
 3. (2 marks) One application of data bias being introduced into ML algorithms is recidivism framework of criminal justice where data from different judges can cause bias. Summarize the key concepts from [Himabindu et al.](#) which comes up with an algorithm that gives an output close to true failure rate without dwelling into data bias. Intuitively explain how the bounds are constructed.
2. (2 marks) Explain with the help of a causal diagram the different unfair paths in the dataset for recidivism in multiple judges setting? How will data bias analysis change in a recidivism setting for multiple judges ? Work along the same lines as [Silvia et al.](#) by using a changed graph.
 3. (2 marks)* Count based bounds given in [Himabindu et al.](#) paper do not consider the fact that bias existed in the dataset. Do we need to tweak the algorithm to account for data bias? Intuitively we know the bounds have to change after accounting for bias. Give a broad level explanation on how to alter the bounds on using ideas from the above question.
-