Assignment #9

Li Liu

Referee report for Athey (2018, forthcoming)

The research question of Athey (2018) is: What is the impact of machine learning on economics? In the introduction part, the author defines the question by elaborating on each part of the paper and several themes of machine learning algorithms. Specifically, the "impact" would be examined by surveying the economics literature that has "new questions", "new approaches to collaboration", or "a change in how involved economists are in the engineering and implementation of policies" (Athey, 2018, p. 2).

The author persuasively answers the research question in a clear structure. Firstly, the author defines machine learning as "algorithms designed to be applied to datasets" and compares it with "casual effect estimation" in the empirical economics literature (Athey, 2018, p. 3-5). Secondly, the author surveys the applications of predictive analytics in policy problems, such as hip replacement operation, jurisdictions, credit scoring, loan repayment, and advertising optimization (Athey, 2018, p. 7-8). Thirdly, the author explains the new research frontier of applying machine learning in causal inference, including "estimating average heterogeneous treatment effects", "estimating optimal policies", "contextual bandit approach", "robustness measure", "difference-in-difference models", "matrix completion approach", and "structural models" (Athey, 2018, p. 10-20). Lastly, the author predicts the trend in many aspects of the economics research community influenced by machine learning (Athey, 2018, p. 21-22).

The methods to answer the research question are appropriate, but the research question might be too broad to be answered completely in one paper. Most of the applications of machine learning in economics mentioned in Athey (2018) are either in the field of business operations and marketing or related to policy-making in crimes, law, healthcare, etc. These areas are just a small subset of the domain of economics research. Athey (2018, p. 23) mentions macroeconomics briefly in "The emerging literature … on "now-casting" in macroeconomics and ML begins to address some, but not all, of these issues". As an expert in microeconomics and econometrics, Dr. Athey might be more familiar with the literature in her fields than other economics topics, such as behavioral economics, financial economics, monetary and fiscal policies, and international trade. Also, many predictions of impact, including "adoption of collaboration tools" and " increase in interdisciplinary work", come from the innovations and advancements of data science community instead of machine learning algorithms alone (Athey, 2018, p.24). As a result, the research question might be more precisely defined as the impact of machine learning and data science on business economics.

The author does an excellent work in reviewing literature in economics, machine learning, statistics, and public policy. The author mentions machine learning could analyze unstructured data, such as images and text, but doesn't provide many applications (Athey, 2018, p.3). Here are two relevant applications that could be surveyed in the paper. Gans, Goldfarb, and Lederman (2017) perform textual analysis to classify sentiments of 4 million tweets and construct variables to measure customers' loyalty to the airline company. Donaldson and Storeygard (2016) show

how machine learning and computer vision could transform the satellite data into useful predictors for economics research. Citing these relevant research articles could help the author better answer the research question.

There are several grammatical and spelling errors in the current draft of Athey (2018). On page 7, "...where the probabiity of selecting…" should be replaced as "probability". On page 10, the attribute "particularly economists and other social scientists" is redundant to describe "social scientists". On page 11, "First" doesn't seem to be necessary for the sentence "First, we can consider…". On page 18, there is a " ' " missing in "...since we dont observe…". On page 20, there is a " ' " missing in "...both users willingness…". On page 22, "...this article has al discussed…" should be "already". On page 26, there should be a word "courses" in "...take coding and data analysis (courses) in college". The references are correctly listed.

Athey (2018, p.1) provides an assessment and prediction framework to study the current and future "impact of machine learning on economics". This method can be extended to study the impact of machine learning on other social science disciplines, such as sociology, political science, psychology, etc. These fields are similar to economics as they focus on causal inference questions, but they are different from economics as they might be less associated with business applications and policy-making. For example, if we want to study the impact of machine learning on sociology, we could start by reviewing the recent sociology literature using machine learning and big data toolkits. Then we contrast it with the literature with traditional approaches, such as social networks and regression analysis. Also, we could predict the future impact by evaluating

the advantages and disadvantages of using the always-on large data, such as Twitter feeds, Facebook posts, Linkedin networks, etc. Finally, we could point out the importance of ethics in sociology research by case studies and further discuss the ethical issues of applying machine learning algorithms in people's daily lives.

Reference

Donaldson, D., Storeygard, A. (2016). The view from above: Applications of satellite data in Economics. *Journal of Economic Perspectives, 30(4)*:171–198.

Gans, J., Goldfarb, A., Lederman, M. (2017). Exit, tweets and loyalty. *Working paper,* University of Toronto, Toronto