Referee Report on Athey (2018)

The Research Questions

The author defines the research question clearly in the abstract of the article. The question that the article aims to answer is what are the early and potential future impact of machine learning on economics research? Overall, the author compellingly answers the research question with appropriate and sufficient methods. The author follows the following logics in her answer.

First of all, the author offers a narrow definition of machine learning: "a field that develops algorithms designed to be applied to datasets, with the main areas of focus being prediction (regression), classification, and clustering or grouping tasks" (Athey, 2018, p.3). She also introduces supervised and unsupervised machine learning methods and their usages briefly. Unsupervised learning is commonly used to categorize items with similar characters, and supervised learning is typically employed to predict an outcome using a set of features (Athey, 2018, p.3).

After introducing the definition of machine learning, the author identifies and compares the major concerns of machine learning and empirical economics. The goal of machine learning is to "achieve goodness of fit in an independent test by minimizing deviation between actual outcomes and predicted outcomes", while that of applied econometrics is to "understand an object and evaluate the impact of covariates" (Athey, 2018, p.4). The author describes machine learning as an algorithm to select the best model maximizing a criterion from many models, which is a data-driven model selection. In machine learning, there is usually "a tradeoff between expressiveness of the model and risk of over-fitting", and multiple approaches were employed to balance them (Athey, 2018, p.4). In the field of economics, the researchers usually "specify one model estimating the full dataset" and "rely on statistical theory to estimate confidence intervals for estimated parameters" (Athey, 2018, p.5). In these study, the researchers focus on "estimated effects rather than goodness of fit of the model" and the "estimate of a causal effect" (Athey, 2018, p.5). The author raises the concern that though researchers only propose one model in a study, they commonly check many models, and they are not systematic or comprehensive in their checking (Athey, 2018, p.5). She proposes that machine learning can "regularize and systematize" the model selecting process, making the process much more effective (Athey, 2018, p.5). However, I found the author's logic a little confusing here. While the author mentions briefly that the application of machine learning in economics might be problematic because of the different goals of machine learning and empirical economics, she fails to elaborate on this issue. By emphasizing on the improvement machine learning can bring to the model searching process, the author does not consider that the criterion of a good model in machine learning probably does not match with that in econometrics. Though the author answers this question later in the article, this first section of the article is confusing. Another concern in applied economics that the author mentions is whether "the assumptions required to identify a causal effect are satisfied", and machine learning cannot address this issue (Athey, 2018, p.6). Machine learning can only "make estimation methods more credible, while maintaining the identifying assumptions" (Athey, 2018, p.6).

The author then reviews some applications of machine learning in economics, especially in the field of policy analysis. Despite the wide range of application of machine learning in policy analysis, there are many concerns about it, including “interpretability of models”, “fairness and nondiscrimination promoted by models”, “stability and robustness regarding variations”, “manipulability of the data”, “computational time”, and “cost of collecting and maintaining the features used in models” (Athey, 2018, p.8-11). The author is optimistic about ways social scientists will define and propose solutions to these concerns (Athey, 2018, p.12).

The author then provides an overview of concerns regarding the expected emerging literature “combining machine learning and causal inference to create new methods that harness the strengths of ML algorithms to solve causal inference problems” (Athey, 2018, p.21). She offers insight in some possible strategies that the paper can employ, such as "average treatment effects", "heterogeneous treatment effects and optimal policies", "contextual bandits", "robustness and supplementary analysis", "panel data and difference-in-difference models", and "factor models and structural models" (Athey, 2018, p.12-21).

Finally, the author describes her prediction of machine learning application in the field of economics as a whole. She proposes that the “combination of machine learning and newly available datasets” can change empirical economics fundamentally with “new questions”, “new approaches to collaboration”, and “change in economists’ involvement in the engineering and implementation of policies” (Athey, 2018, p.23-26).

Literatures

Overall, the author establishes the idea of the article well with many literatures in different areas including machine learning, economics, and statistics. However, after careful inspection, I notice that there is an uneven distribution of literatures throughout the paper.

In the first part of the article, the author well defines the definition of machine learning, discusses the key concerns of economics and machine learning, and proposes ways to combine them regarding these concerns. Though the content follows a great logic, the author falls short of providing enough literatures to support her idea. In my opinion, the author can include some literature supporting her narrow definition of machine learning, the idea that machine learning is “useful as an intermediate step in empirical work in economics” , the bad practice of researchers checking several models without a systematic and comprehensive approach, the focus of economics on causal effect estimation, and the concern of economics about assumption (Athey, 2018, p.5). I believe the author is able to find articles supporting the founding concepts of her arguments. The author does refer to many early applications of machine learning in studies in different areas, such as McFadden et al (1972), Laffont et al. (1995), Athey et al. (2011), Athey et al. (2013), and Athey and Haile (2007).

In the section on the application of prediction methods in policy analysis, the author provides sufficient amount of literatures to illustrate the examples of successful applications. In contrast, the author does not justify the “research questions arise when prediction methods are taken into policy applications” (Athey, 2018, p. 8). The author only cites her own previous study (Athey, 2017) and four other studies including Imbens and Rubin (2015), Yeomans et al. (2016), Kleinberg et al. (2016), and Bjorkegren and Grissen (2015) to arguing the five issues for the application of prediction methods in policy studies (Athey, 2018. p.8-11).

In the section on strategies combining machine learning and causal inference, the author uses a lot more literatures. Though using many literatures is good for supporting her arguments, comparing to the scarce literatures used in other sections, the author may use some repetitive and unnecessary literatures in this section.

Finally, in the last section predicting future of machine learning in social science research, the author uses sufficient and appropriate literatures.

Extension

While the author does answer the research question appropriately and sufficiently, there is still space for improvement in discussing the motivation of incorporating machine learning in the area of economics or social science in general.

As I mentioned, the author fails to elaborate clearly how machine learning could address the major concerns in empirical economics. Moreover, besides providing many examples of successful applications of machine learning in the area of policy analysis, the author does not mention the motivation or clearly state the advantage of using machine learning comparing to using traditional approaches. My suggestion is to use more literatures in discussing benefits of using machine learning and the motivations for applications in previous studies and for further exploration in the future.

Grammatical, spelling, or style erros

There are lots of grammar mistakes, especially spelling mistakes, in the article, and they are shown in the list below. The errors are very obvious, so I do not include the corrected forms.

1. "<u>the the</u> emerging literature in econometrics" (Athey, 2018, p.1)
2. " <u>a collections</u> of subfields of computer science" (Athey, 2018, p.2)
3. "This approach <u>constrasts</u> with economics" (Athey, 2018, p.2)
4. "a human watches <u>the the</u> largest group" (Athey, 2018, p.3)
5. " <u>becaues</u> the affect both the optimal price set by the firm" (Athey, 2018, p5)
6. "they <u>can not</u> be rejected by looking at the data" (Athey, 2018, p.6)
7. " a <u>probability</u> distribution over bidder values" (Athey, 2018, p.6)
8. "subtle relationships <u>bewteen</u> X and Y" (Athey, 2018, p9)
9. "pianos are not a <u>fundamentnal</u> feature of cats" (Athey, 2018, p.9)
10. "consumers may be able to <u>maniplate</u> the data " (Athey, 2018, p.9)
11. "changes prices at a given point in <u>timeat</u> doing" (Athey, 2018, p.11)
12. "experiment-like variation in <u>pricesacrificing</u> predictive accuracy" (Athey, 2018, p.11)
13. "combination of identification <u>strategty</u> and problem of interest" (Athey, 2018, p.11)
14. "the average effect if a treatment <u>were</u> applied to" (Athey, 2018, p.8)

15. “rather than hetereogeneity” (Athey, 2018, p.15)
16. “ asksthe question” (Athey, 2018, p.16)
17. “learning for the future indirectly indirectly” (Athey, 2018, p.17)
18. “whereby the analyst assses whether ” (Athey, 2018, p.17)
19. “since we dont observe” (Athey, 2018, p.18)
20. “serves at a potential source of criticism of a paper” (Athey, 2018, p.22)
21. “more directly useful quantiatively” (Athey, 2018, p.23)
22. “facilitates measurement and caual inference” (Athey, 2018, p.24)
23. “emphasis on documenation and reproducibility” (Athey, 2018, p.24)
24. “individual researcheres” (Athey, 2018, p.24)
25. “delivered digitially” (Athey, 2018, p.24)
26. “Google’s web site” (Athey, 2018, p.25)
27. “arrive at a web site” (Athey, 2018, p.25)
28. “precison of estimates” (Athey, 2018, p.25)
29. “There have already been a wide range of applications” (Athey, 2018, p.14)
30. “The fact that the exploration benefits the future through a model of how contexts relates to outcomes changes the problem” (Athey, 2018, p.17)

References

1. S. Athey and P. A. Haile. Nonparametric approaches to auctions. Handbook of econometrics, 6: 3847–3965, 2007.

2. S. Athey, J. Levin, and E. Seira. Comparing open and sealed bid auctions: Evidence from timber auctions. The Quarterly Journal of Economics, 126(1):207–257, 2011.

3. S. Athey, D. Coey, and J. Levin. Set-asides and subsidies in auctions. American Economic Journal: Microeconomics, 5(1):1–27, 2013.

4. S. Athey, G. Imbens, T. Pham, and S. Wager. Estimating average treatment effects: Supplementary analyses and remaining challenges. American Economic Review, 107(5):278–81, 2017c.

5. S. Athey. The impact of machine learning on economics. University of Chicago Press. 2018.

6. D. Bjorkegren and D. Grissen. Behavior revealed in mobile phone usage predicts loan repayment. 2015.

7. G. W. Imbens and D. B. Rubin. Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press, 2015.

8. J. Kleinberg, J. Ludwig, S. Mullainathan, and Z. Obermeyer. Prediction policy problems. American Economic Review, 105(5):491–95, 2015.

9. J.-J. Laffont, H. Ossard, and Q. Vuong. Econometrics of first-price auctions. Econometrica: Journal of the Econometric Society, pages 953–980, 1995.

10. D. McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.

11. M. Yeomans, A. K. Shah, and J. Kleinberg. Making Sense of Recommendations. 2016. URL http://goo.gl/8BjhMN.