▬▬▬▬▬▬   BLOG POST

Machine learning and economics

Machine learning (ML), together with artificial intelligence (AI), is a hot topic. Economists have been looking into machine learning applications not only to obtain better prediction, but also for policy targeting. We review some of the contributions.

BY: SILVIA MERLER   **DATE:** NOVEMBER 29, 2018
TOPIC: INNOVATION & COMPETITION POLICY

(🐦)   (f)   (in)   (✉)

Writing on PWC blog last year, **Hugh Dance and John Hawksworth** discussed what machine learning (ML) could do for economics in the future. One aspect is that of prediction vs causal inference. Standard econometric models are well suited to understanding causal relationships between different aspects of the economy, but when it comes to prediction they tend to "over-fit" samples and sometimes generalise poorly to new, unseen data.

By focusing on prediction problems, machine learning models can instead minimise forecasting error by trading off bias and variance. Moreover, while econometric models are best kept relatively simple and easy to interpret, ML methods are capable of handling huge amounts of data, often without sacrificing interpretation.

Susan Athey provides an assessment of the early contributions of ML to economics, as well as predictions about its future contributions. At the outset, the paper highlights that ML does not add much to questions about identification, which are of concern when the object of interest, e.g. a causal effect, can be estimated with infinite data. Rather, ML yields great improvements when the goal is semi-parametric estimation or when there are a large number of

covariates relative to the number of observations.

The second theme is that a key advantage of ML is that it views empirical analysis as "algorithms" that estimate and compare many alternative models. This approach contrasts with economics, where in principle the researcher picks a model based on principles and estimates it once. The third theme deals with the "outsourcing" of model selection to algorithm. While it manages the "simple" problems fairly well, it is not well suited for the problems of greatest interest for empirical researchers in economics, such as causal inference, where there is typically no unbiased estimate of the ground truth available for comparison. Finally, the paper notes that the algorithms also have to be modified to provide valid confidence intervals for estimated effects when the data is used to select the model.

Athey thus thinks that using ML can provide the best of both worlds: the model selection is data-driven, systematic, and considers a wide range of models; all the while, the model selection process is fully documented and confidence intervals take the entire algorithm into account. She also expects the combination of ML and newly available datasets to change economics in fundamental ways, ranging from new questions and new approaches to collaboration (larger teams and interdisciplinary interaction), to a change in how involved economists are in the engineering and implementation of policies.

David McKenzie writes on the World Bank blog that ML can be used for development interventions and impact evaluations, in measuring outcomes and targeting treatments, measuring heterogeneity, and taking care of confounders.

One of the biggest use cases currently seems to be in getting basic measurements in countries with numerous gaps in the basic statistics – and machine learning has been applied to this at both the macro- and micro-level. But scholars are increasingly looking into whether ML could be useful also in targeting interventions, i.e. in deciding when and where/for whom to intervene.

McKenzie, however, points to several unanswered challenges, such as the question of the gold standard for evaluating these methods. Supervised ML requires a labelled training dataset and a metric for evaluating performance, but the very lack of data that these approaches are trying to solve also makes it hard to evaluate. Second, there are concerns about how stable many of the predicted relationships are, and about the behavioural responses that could affect the reliability of ML for treatment selection. Lastly, McKenzie also points to several ethics, privacy and fairness issues that could come into play.

Monica Andini, Emanuele Ciani, Guido de Blasio, Alessio D'Ignazio take up the question of policy targeting with ML in a recent VoxEU article and two papers. They present two examples of how to employ ML to target those groups that could plausibly gain more from the policy.

One example considers a tax rebate scheme introduced in Italy in 2014 with the purpose of boosting household consumption. The Italian government opted for a coarse targeting rule and provided the rebate only to employees with annual income between €8,145 and €26,000. Given the policy objective, an alternative could have been to target consumption-constrained households who are supposed to consume more out of the bonus. Applying ML to two household survey waves, the authors implement the second strategy. An additional econometric analysis suggests that this version of targeting would have been better, because the effect of the rebate is estimated to be positive only for the consumption-constrained households.

In the second application, **Andini et al.** focus on the "prediction policy problem" of assigning public credit guarantees to firms. In principle, guarantee schemes should target firms that are both creditworthy and rationed in their access to credit. In practice, existing guarantee schemes are usually based on naïve rules that exclude borrowers with low creditworthiness.

Aldini et al. propose an alternative assignment mechanism based on ML predictions, in which both creditworthiness and credit rationing are explicitly addressed. A simple comparison of the growth rate of disbursed bank loans in the years following the provision of the guarantee shows –

confirmed by a regression discontinuity design
– that the ML-targeted group performed better.

Republishing and referencing

Bruegel considers itself a public good and
takes no institutional standpoint. Anyone is
free to republish and/or quote this post
without prior consent. Please provide a full
reference, clearly stating Bruegel and the
relevant author as the source, and include a
prominent hyperlink to the original post.

🏷 AI   🏷 ARTIFICIAL INTELLIGENCE

🏷 ECONOMIC BLOGS REVIEW   🏷 MACHINE LEARNING