Instructions for the Final Project

Fall 2021

Final Project

- Students will be asked to find data from a published economics paper (in Chinese or English), re-do the econometric analysis using one of the ML methods taught in the course, discuss the differences and finally write the report as a short academic article (in Chinese or English).
- The focus should be on estimation of causal effects and not prediction/forecasting.
- Our goal is to evaluate the usefulness of ML tools for estimation of causal effects by comparing their performance with that of traditional econometric methods.

Assessment Criteria

- Motivation of the chosen research question.
- The quality of presentation of the econometric model and formulation of the research question in terms of the model.
- Motivation for using ML techniques.
- Whether you apply the ML techniques correctly and thoroughly.
- Whether you Interpret and analyze the results thoroughly.
- Quality of your academic writing.

The Toolbox

- Adaptive LASSO: one-step model selection and estimation for the classical linear regression model using LASSO.
- Double LASSO: a more sophisticated technique in the context of sparse high-dimensional linear regression model with many controls.
- LASSO for IV models: (1) many IVs and a small number of controls; (2) a small number of IVs and many controls; (3) many IVs and many controls.
- Causal forests: estimating the effect of a binary variable on the outcome using a model that is different from the linear regression model.

Format

- You are encouraged to use LyX or Latex. But an article prepared by MS Word is still acceptable.
- You should also hand in your R code in the form of an RMarkdown report that generates your outputs.
- In the RMarkdown file, please briefly describe what you do in each step.

Structure of the Article

1. Motivation and research question.
2. Data description.
3. Describe the econometric model.
4. Describe the results and conclusion produced by traditional methods.
5. Provide your results using ML techniques.
6. Compare the two sets of results. Any similarity or difference?
7. Provide your discussion and make your conclusion.

Useful Resources

- Find Economic Articles with Data: http://econ.mathematik.uni-ulm.de:3200/ejd/
- LyX: https://www.lyx.org/
- Library genesis: http://libgen.rs/
- Sci-Hub: https://www.sci-hub.ren/