

Crossing the Bridge

From Econometrics to Machine Learning

Philippe Goulet Coulombe

goulet_coulombe.philippe@uqam.ca

Université du Québec à Montréal

January 11, 2023



Different Objectives (in principle)

- Classical econometrics, except for the forecasting literature : $\hat{\beta} \rightarrow_p \beta$
- No way to empirically verify how close to the objective we are, typically works with "arguments", like exogeneity of some IV
- **Result:** causal analysis
- ML (most of it) : $\min \sigma_{\epsilon}^2$, aka obtain the conditional mean $E(y_i|X_i)$. But, also, there is nothing sacred about the mean.
- Can be evaluated by looking at how well a model does on fresh data ($\{y_i, X_i\}$ pairs that were not used to estimate \hat{f} , or $\hat{\beta}$ if the model's linear)
- **Result:** prediction
- However, things are not so black and white anymore...

Machine Learning Principles inherited from... Learning

- Some tasks are more complex than others
- Practice make perfect
- More practice is needed for more complex tasks,
- When faced with a task, one can either learn or imitate. While the results of both are indistinguishable on the task then were trained on, imitation will fail for new (yet related) task. Think about exams in school.
- Powerful Learning *algorithms* typically start by learning, then imitate. "Tuning" is required to make sure only the former occurs.

Pieces of Empirical Wisdom to Avoid Doing Stupid Things

- No harm will ever come in looking at the data first.
- Always do the simplest thing first
- If things are looking too good (especially with traditional economic data), there is likely a problem with the code
- Always compare your new exciting results to the simplest thing.
Examples:
 - Macro: AR
 - Finance (returns): RW
- Do not forget basic econometric wisdom. Examples:
 - ML works *much* better with stationary data, so did traditional frequentist inference in time series
 - Heteroscedasticity may need to be addressed (for different reasons)

A Sales' Pitch

- ML makes you understand econometrics and statistics better (different viewpoint on many common themes)
- Learning ML will allow you to dialog with many people within many fields outside the tribe of economists, and perhaps even claim the title of "data scientist"
- Industry likes ML
- We have probably only seen the tip of the iceberg