Thesis topic research

Lingjie

April 5, 2022

Contents

1 Overview

Applied research on solving economics issues with Machine Learning techniques. The thesis paper will focus on the framework and use case (HOW do we apply machine learning techniques and WHAT benefits we can gain).

Proposed structure

1. Research on existing framework and algorithms
    [1.1] Explaining algorithms
    [1.2] Comparing performance
2. Applying framework and algorithms to illustrate the benefit it brings to economic field
    [2.1] classic data (with known ground truth or baseline for comparison)
    [2.2] new dataset to illustrate usefulness

Relevant keyword: Computational Economics, machine learning, economic + *topic keywords*

Search words: machine learning on economics

1.1 Some questions regarding thesis paper

How in-depth should an undergraduate thesis be? → what is the expectation? I saw some thesis simply compare off the shelve ML models without hyper-parameter tuning.

2 Possible topics

2.1 ML + Causal inference

Motivation: Social science is unique in its causal inference study. Other data domains such as statistics and computer science focus on prediction instead causation.

I hope to explore research topics on applying machine learning (and other techniques) to estimate heterogeneous causal effects.

Concerns: I have no experience in causal inference and I'm uncertain how 'impactful' a causal paper needs to be. I have passed causal inference course in Coursera and read some discussions on causal inference. I will be taking EC4305 Applied Econometrics next semester to have a formal learning.

2.1.1 Relevant papers

Framework

- Machine Learning for Estimating Heterogeneous Causal Effects
- Estimating treatment effect heterogeneity in randomized program evaluation
- Modelling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees
- Machine Learning for Causal Inference: On the Use of Cross-fit Estimators
- a list of research papers on causal ML

Sample use case

- Machine learning for causal inference in Biostatistics

Open-source library

- Uber causal ML
- Microsoft causal ML and Microsoft Causality and Machine Learning
- paid course on causal ML

2.2 ML + Forecasting

Motivation: Machine learning (and deep learning) introduces more toolkits for economists to apply on forecasting problems.

Perhaps we can explore the benefit ML techniques brings to a specific domain or problem.

Concerns: Although I'm most confident with forecasting models, I'm not sure if writing a thesis on this topic will be interesting. Furthermore, I'm not sure what is the value added to economic field with another model fitting.

2.2.1 Relevant papers

Sample use case

- Machine Learning in Economics and Finance

3 Possible dataset

3.1 Open data

Possible open sourced dataset (either through web scraping or direct download)

- Social media website (e.g. Twitter, Facebook, Linkedin)
- Google mobility data
- Financial data

  Bitcoin/stock market

  Financial news (NLP)
- Macro-economic indicators

  GDP/employment data