

まとめ

機械学習

川田恵介 (keisukekawata@iss.u-tokyo.ac.jp)

Table of contents

| | | |
|---|--------------------|---|
| 1 | まとめ | 1 |
| 2 | 機械学習 (教師付き学習) と統計学 | 1 |
| 3 | 代表的なゴール | 2 |
| 4 | 別のゴール: 因果推論 | 2 |
| 5 | イメージ図 | 2 |
| 6 | 復習 | 3 |

1 まとめ

- 将来の学習に向けた整理
 - 本講義は大きく、機械学習を用いた予測モデル推定 と予測モデルの比較分析への活用、を議論

2 機械学習 (教師付き学習) と統計学

- 本講義では、教師付き学習の教科書ではよく紹介されるが、統計学の伝統的な教科書ではあまり紹介されない手法 (決定木、Random Forest) を主に紹介
 - 共通して紹介される概念/手法は、非常に多い (過剰適合/過学習、OLS, Logit/Probit 等)
- (私見) データ分析をこれから本格的に学ぶ人は、区別しない方が実践的
 - “歴史的に異なるルーツを持つが、現状では融合”
- むしろ異なるゴールをしっかり意識

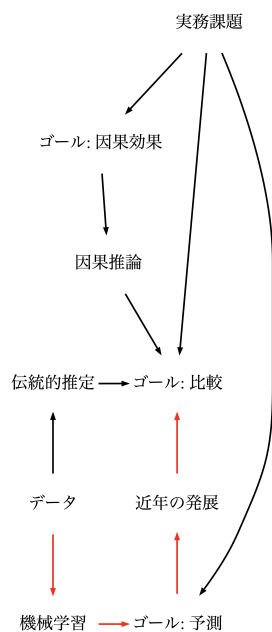
3 代表的なゴール

- Y の予測 (Prediction): 「機械学習」と特に親和的
- 母集団の特徴の要約/説明 (Description/Explanation): 「統計学的手法」と特に親和的
- 現状では融合が進み、予測を比較に補助的に活用するハイブリットな方法が発展

4 別のゴール: 因果推論

- 踏み込んだ説明: 母集団の因果的特徴を明らかにする
 - グループの比較や予測ではなく、“介入” が引き起こす“変化” を明らかにしたい
 - ほぼほぼ議論は、どのような比較分析を行えば、因果効果を明らかにできるのかを議論
- 因果推論に必要な比較分析を行うためのツールとして、機械学習も活用可能: 極めて活発な研究分野
 - [ノーベル記念経済学賞](#)
 - [Rousseeuw Prize](#) (統計学版ノーベル賞)
 - [ALICE project](#)

5 イメージ図



6 復習

- [CausalML](#) (1, 3, 4, 9, 10, 14,15 章)
 - 他にも因果推論の内容をカバー