

宿題 3

機械学習

川田恵介

2025-11-29

提出方法: Posit cloud からダウンロードした ScriptFile を、Musashi 3S をアップロードする

提出物: スクリプトファイルに分析コードと、記述問題の回答をコメントの形で書き込む

提出期限: 2025 年 12 月 31 日

以下を講義内容を踏まえて、答えよ

Q1

- Random Forest 法に比べて、**OLS 法が持つ利点**を答えよ

Q2

- 今、武藏大学の経済学部と社会学部生を対象に、アンケートを行ったとします。このアンケートをもとに、1ヶ月に衣服に使う平均金額を経済学部と社会学部で比較する分析を試みます。またこの際に、学生の出身都道府県と年齢をバランスさせたいとします。機械学習による予測モデル推定を活用しながら、この比較分析を行う方法を可能な限り詳しく説明せよ。

Q3

講義で用いたデータを使って、 $Y = \text{駅からの距離}$ 、を予測するモデルを Stacking 法で推定し、サンプル分割法を用いて評価せよ

- ヒント: SuperLearner を使う

Q4

$Y = \text{築年数}$ 、 $D = \text{改装の有無}$ 、の関係性を予測モデルを活用しながら推定せよ。その際に、 $X = \text{取引価格}$ 、駅からの距離、をバランスせよ

- ヒント: ddml を使う