

「社会科学のための機械学習入門：コンセプトから R による実装まで」

日時：2020年3月6日（金）10:30～17:00

場所：東京大学（本郷キャンパス）赤門総合研究棟5階・センター会議室（549号室）

料金：一般 5,000円、学生 2,500円

講師：川田恵介（東京大学）

定員：25名

■本コースの内容

「機械学習の社会科学への応用」について基礎から学ぶコースです。機械学習は、データを“柔軟”に用いることで、性能の良い予測モデルを構築することを可能にします。標準的な toolkit の一つとして学部/大学院レベルの計量経済学の教科書でも紹介されるなど、社会科学においても急速に活用が拡大しております。

本コースでは、機械学習の基礎から社会科学の伝統的な関心である因果推論/比較研究への応用まで、高度な数学を用いずに紹介します。現状でデータを扱っている研究者・分析者のみならず、これからデータ分析をやりたい院生・学生の参加も歓迎いたします。

さらに無料の統計ソフトである R と SSJ データアーカイブに寄託されている個票データを用いて、実際に自分の手で統計的因果推論を用いた実証研究を体験してもらいます。こちらについても基礎的な内容から始めますが、R や因果推論に初めて触れられる方は、3月5日に開催される「回帰分析と傾向スコアを用いた因果推論」コースも併せて受講されることをお勧めします。

■次のような方におすすめです

- ・機械学習、とくに社会科学研究への応用に関心がある方
- ・「機械学習/AI を用いたデータの解析結果」を批判的に検討したい方
- ・新しい統計ソフトの習熟に忌避感の無い方

■注意事項

- ・SSJ データアーカイブのデータを利用するため、大学または公的研究機関所属の研究者・学生（学部生も可）、SSJ データアーカイブへデータを寄託されている民間研究機関の方のみを対象とします。
- ・R および Rstudio がインストールされたノートパソコンをご持参ください。インストールに問題がある場合は、当日 10:00 までに会場にお越しください。

■本コースの日程

「過剰適合問題」、「データ分割」、「回帰系アルゴリズム」、「決定/回帰木系アルゴリズム」、
「コントロール変数法への応用」、「効果の異質性分析への応用」

* 進度によって、内容が若干変わることがあります。