

授業概要

データサイエンスには「プロセスとしてのデータサイエンス」と「学問としてのデータサイエンス」があります。この講義では「プロセスとしてのデータサイエンス」とはどのようなものであるかを講義し、その分析手法の内の代表的なものを解説し、実習いたします。どれも古典的基本的な手法です。AIとデータサイエンスの関係や最先端の手法の話は最終回に少し触れるだけになります。

授業計画

第 1 回	オリエンテーション（データサイエンスとは）
第 2 回	Rの環境と基本操作
第 3 回	Rのファイル操作
第 4 回	線形回帰分析①（理論）
第 5 回	線形回帰分析②（演習）
第 6 回	アソシエーション分析①（理論）
第 7 回	アソシエーション分析②（演習）
第 8 回	決定木分析①（基本原理）
第 9 回	決定木分析②（最適木）
第 10 回	決定木分析③（演習）
第 11 回	クラスター分析①（階層的クラスター分析・演習）
第 12 回	クラスター分析②（k-means・演習）
第 13 回	ニューラルネットワーク①（理論）
第 14 回	ニューラルネットワーク②（演習）
第 15 回	ニューラルネットワーク③（さまざまなモデル）・データサイエンスの動向
第 16 回	定期試験

到達目標

データサイエンスとは何かを理解する。

古典的基本的な分析手法を理解し、課題設定・データ分析のデザイン・データ収集・分析・分析結果からの価値創造の一連の過程を自分で実行できるようになる。

AIとデータサイエンスの関係を理解する。

履修上の注意

「経済統計学」などの統計系科目を履修済みか履修中であることが望ましい。

また、全学共通科目の「数学（線形代数基礎）」「数学（解析基礎）」の履修を強く勧めます。

Rというソフトウェアを利用して実習を行うため、プログラム言語 Python に関する知識は特に必要ではありません。ただし、Python を用いた分析についての紹介はいたします。

予習・復習

予習：教科書で次回の内容を予習し講義に臨んでください。

復習：授業内容を整理して理解を深めてください。

評価方法

第 7 回終了後の課題（40 点）と期末試験（60 点）の合計 100 点のうち 50 点以上を取れば合格となります。ただし、出席回数が 10 回に満たない人は成績評価できませんので注意してください。

テキスト

- ・教科書名：Rによるデータサイエンス 第2版
- ・著者名：金 明哲
- ・出版社名：森北出版株式会社
- ・ISBN：978-4-627-09602-8