

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説明し、その分析手法の内の代表的なものを解説し、実習します。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第1回と第15回に説明します。この科目はデータサイエンス科目群に属します。

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）、Rの環境と基本操作
第2回	Rのオブジェクトとファイル操作
第3回	回帰分析①（理論・線形単回帰）
第4回	回帰分析②（線形重回帰・非線形回帰）
第5回	アソシエーション分析①（理論）
第6回	アソシエーション分析②（演習）
第7回	決定木分析①（基本原理）
第8回	決定木分析②（最適木）
第9回	決定木分析③（演習）
第10回	クラスター分析①（階層的クラスター分析・演習）
第11回	クラスター分析②（k-means・演習）
第12回	ニューラルネットワーク①（理論）
第13回	ニューラルネットワーク②（単一中間層モデル、深層ニューラルネットワークの準備）
第14回	ニューラルネットワーク③（深層ニューラルネットワーク）
第15回	ニューラルネットワーク④（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できる。
- ・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。
- ・AI とデータサイエンスの関係を理解できる。

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6回終了後の課題（40点）と期末試験（60点）の合計100点のうち51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