科目ナンバリング	U-LAS11 10007 LJ55						
授業科目名 <英訳>	データ分析基礎 Basic Data Analysis			担当者所属 職名・氏名	国際高等教育院 教授　　原　尚幸		
群	自然科学科目群			分野(分類)	データ科学(基礎)	使用言語	日本語
旧群	B群	単位数	2単位	週コマ数	1コマ	授業形態	講義（対面授業科目）
開講年度・ 開講期	2026・前期	曜時限	火5	配当学年	全回生	対象学生	全学向

[授業の概要・目的]

本講義は、近年のビッグデータ解析において重要な役割を果たしている機械学習・AIの基礎理論である多変量解析の諸手法を理論・実践の両面から学習し、多次元データが与えられたときに、適切な手法を選んで能動的にデータ分析を行い、結果を正しく解釈できるようになることを目的とする。
具体的には、回帰分析、クラスター分析、判別分析、主成分分析、対応分析、正準相関分析などの手法について、実用例とともに解説を行う。また、近年の機械学習・AIへの接続として、スパース法やカーネル法の初歩的な考え方についても解説を行う。
線形代数・（高校数III程度の）微積分のごく基本的な知識があることが望ましい。
また、統計入門は受講済であることが望ましい。

本講義の単位（2単位）を修得することで、文部科学省が定める数理・データサイエンス・AI教育プログラム応用基礎レベル(MDASH Advanced Literacy)修了証の取得が可能である。
修了証取得の手続きについては、講義内で担当教員より指示がある。

[到達目標]

1. 近年における多次元データの重要性について理解をする。
2. 多変量解析の諸手法について概念、特に「次元縮約の概念」と、その使いどころを理解し、多次元のデータ分析に応用できるようになる。
3. 統計ソフトRを使いこなす知識を身につけ、実際に簡単なデータ分析を行えるようになる。
4. 近年の機械学習における多変量解析の展開を理解する。

[授業計画と内容]

授業回数はフィードバックを含め全15回とする。
開講にあたっては、受講生の所属するキャンパスの配置や受講形態にも配慮し、一部メディア授業を取り入れることがある。

・ガイダンス（1回）
授業の概要を確認するとともに、本講義で主として学習する多変量解析の考え方を概説する。

・データリテラシー１（1回）
近年のAI技術の利活用と課題について解説を行う。

・回帰分析（3回）
機械学習・AIを用いた予測の基礎技術である回帰分析について学習する。最小二乗法を用いたデータ分析の方法について詳しく解説を行った後、機械学習への接続としてスパース法や非線形回帰などについても実践的に学ぶ。

・クラスター分析（2回）
教師なし学習の代表的な技術であるクラスター分析の考え方について学習する。階層的クラスター

データ分析基礎(2)へ続く↓↓↓

データ分析基礎(2)

分析・非階層的クラスター分析に加え、多次元尺度構成法について学習する。

・主成分分析（2回）
多変量解析の基礎である主成分分析について学習する。基本的な線形代数を用いて「データ縮約」の考え方を学ぶ。

・判別分析（2回）
古典的な正準判別分析の初概念を線形代数を用いて解説する。また、サポートベクターマシンのような機械学習的な手法についても概説する。

・対応分析（2回）
質的データの多変量解析手法である対応分析と結果の可視化の手法について概説を行う。

データリテラシー（２）（1回）
データ駆動型社会、Society 5.0、ビッグデータ、プライバシー保護、個人情報の取り扱いについて解説する。

・フィードバック（1回）

[履修要件]

全学共通科目の統計入門、あるいは数理統計で学習した知識を一部前提とするので、これらのいずれかを受講済であることを推奨する。

[成績評価の方法・観点]

定期試験、宿題（平常点）を総合的に評価する。

・期末試験・期末レポート（配点50点）
講義で解説したデータ分析の基本的な原理や理論を理解できているかを評価する。

・宿題評価（配点50点）
講義の内容を確認するためのレポートや宿題を複数回課すことで理解度を確認し成績評価に加える。宿題に対する評価を平常点とする。

[教科書]

使用しない

[参考書等]

（参考書）
京都大学　データ科学イノベーション教育研究センター『講義実録　統計入門』（現代図書, 2023）
ISBN:978-4-434-31857-3

[授業外学修（予習・復習）等]

多変量解析ではベクトル・行列の演算を頻繁に用いる。
ベクトル・行列については授業中にも解説するが、扱いに慣れていない場合は予習あるいは復習をすることが望ましい。
また、厳密な数学的議論など授業中に省略した事項について、各自で学習することを期待する。

また、成績評価において課題、レポートのウェイトが高い（70%）。課題やレポートの負荷はそれほど高くはないので、確実に対応をすること。

データ分析基礎(3)へ続く↓↓↓

データ分析基礎(3)

[その他（オフィスアワー等）]

授業時間外で質問がある場合には、下記のアドレスにメールで連絡すること。
原　尚幸（はらひさゆき）hara.hisayuki.8k@kyoto-u.ac.jp

[主要授業科目（学部・学科名）]