

科目名	データ駆動科学B
講義題目(テーマ)	機械学習概論と判別・クラスター分析
担当教員	統計数理研究所 日野英逸
年度・学期	2023年 集中
単位数	1

学修成果とその割合	
1.高度な専門的知識・技能及び研究力	70
2.学際的領域を理解できる深奥な教養力	30
3.グローバルな視野と行動力	0
4.地域社会を牽引するリーダー力	0
その他	0

使用言語	「日本語」による授業
教科書・資料の言語	「日本語」のテキスト
実務経験を活かした授業	非該当
授業の形態	講義
対面・遠隔の別	遠隔形式
授業の方法	Zoomを用いた遠隔授業と、オンデマンド受講

授業の目的	機械学習の全体像と、データの特徴量に基づくデータ判別と、クラスター解析法について理解する
授業の概要	以下の事柄について講義を行う。 (1). 機械学習概論、(2).判別分析、(3).ROC解析、(4).サポートベクターマシン、(5).クラスタ分析の基礎、(6).種々のクラスタリング手法、(7).モデルに基づくクラスタリングとクラスタリング結果の評価、(8).情報論的クラスタリング、(9)アンサンブル法

学修目標	
A水準 (到達すれば「優」に相当)	(1). 機械学習概論、(2).判別分析、(3).ROC解析、(4).サポートベクターマシン、(5).クラスタ分析の基礎、(6).種々のクラスタリング手法、(7).モデルに基づくクラスタリングとクラスタリング結果の評価、(8).情報論的クラスタリングについて、十分に理解し、講義内容を他人に正確に説明できる。
C水準 (到達すれば「可」に相当)	(1). 機械学習概論、(2).判別分析、(3).ROC解析、(4).サポートベクターマシン、(5).クラスタ分析の基礎、(6).種々のクラスタリング手法、(7).モデルに基づくクラスタリングとクラスタリング結果の評価、(8).情報論的クラスタリングについて、概ね理解し、講義内容の要点をまとめることができる。
評価方法・基準	Moodleで提出されたレポートの到達度(100%) から評価する。

各回の授業内容		
回	授業テーマ (5文字以上100文字以内)	内容概略(10文字以上200文字以内)
1	機械学習概論	機械学習の諸技術を広く紹介する
2	判別分析1	Fisherの判別分析を中心に判別分析の基礎を説明する
3	判別分析2	ROC解析による判別結果の評価方法とロジスティック回帰分析を説明する

4	判別分析3	サポートベクターマシンの説明をする
5	判別分析4	アンサンブル法と機械学習モデルの解釈性について説明する
6	クラスタ分析1	クラスタ分析の代表的手法を説明する
7	クラスタ分析2	クラスタ分析結果の評価方法とモデルに基づくクラスタリング方法を紹介する
8	クラスタ分析3	情報論的なクラスタリング手法を紹介する

授業外学修時間の目安	本科目は、45時間の学修が必要な内容で構成されている。授業は16時間分（2h×8コマ）となるため、29時間分相当の事前・事後学修（課題等含む）が、授業の理解を深めるために必要となる。
------------	---

キーワード	機械学習，統計科学，判別分析，クラスタ分析，アンサンブル法，解釈性
テキスト	授業の際に資料を配布する。
参考文献	「パターン認識と機械学習，丸善，C. Bishop」 「パターン認識 (Rで学ぶデータサイエンス 5)」

オフィス・アワー	データ駆動型社会を担う人材育成プログラム事務室に連絡をとること
担当教員への連絡方法	データ駆動型社会を担う人材育成プログラム事務室に連絡をとること
担当教員からのメッセージ	はじめに機械学習で扱う様々な問題設定を紹介し、その後、統計的観点から、判別分析とクラスタ分析の基礎を説明します。Rによる諸手法の実行例を交えて説明することで、受講者のみなさんが数理的背景を理解しつつ実問題に適用することができるようになることを目指します。