

JISA D アカデミー プログラム (データサイエンス)

【講義】

No.	日程	講義名	講師	講義内容
1	2021年12月8日(水) 19:00~20:30	データサイエンス入門	村田 剛志 教授 東京工業大学 情報理工学院 情報工学系 教授	<ul style="list-style-type: none"> データサイエンスとはどのような技術なのか 統計学や機械学習の技術との関連 データ解析を行う時に陥りやすい注意点
2	2021年12月15日(水) 19:00~20:30	データ分析の基礎(統計学)	榊島 祥介 教授 東京工業大学 数理・計算科学系(特定教授) 東京大学 大学院理学系研究科附属知の物理学研究センター 教授	<ul style="list-style-type: none"> 大量のデータ群の特徴を抽出する「記述統計」 母集団から標本抽出したデータ群から母集団の特徴を抽出する「推測統計」 主観確率を扱う「ベイズ統計」
3	2021年12月22日(水) 19:00~20:30	分類とモデル評価	榊島 祥介 教授 東京工業大学 数理・計算科学系(特定教授) 東京大学 大学院理学系研究科附属知の物理学研究センター 教授	<ul style="list-style-type: none"> 分類モデル(決定木)の構築 決定木を使った分類の評価基準
4	2022年1月12日(水) 19:00~20:30	主成分分析	関嶋 政和 准教授 東京工業大学 情報理工学院 准教授	<ul style="list-style-type: none"> 多次元の大量データから情報をできるだけ保存しながら情報を縮約する手法 主成分の求め方 分散、共分散、相関係数と主成分の関係
5	2022年1月19日(水) 19:00~20:30	クラスタリング	関嶋 政和 准教授 東京工業大学 情報理工学院 准教授	<ul style="list-style-type: none"> 多次元のデータ間の類似度の定義 大量の観測データを類似度を元にグループ分けするトップダウンの手法とボトムアップの手法
6	2022年1月26日(水) 19:00~20:30	次元圧縮	金森 敬文 教授 東京工業大学 情報理工学院 数理・計算科学系 教授	<ul style="list-style-type: none"> 高次元データから低次元データへの情報圧縮方法 全体的な分布を保存する主成分分析、ペア間の関係を保存する多次元尺度構成法、 二つのデータセット間の共通の構造を保存する正準相関分析
7	2022年2月2日(水) 19:00~20:30	アンサンブル学習	金森 敬文 教授 東京工業大学 情報理工学院 数理・計算科学系 教授	<ul style="list-style-type: none"> 複数の決定木を利用し予測精度を高める方法の学習 多くの決定木の多数決や平均をとるバギング データと着目変数の選択をランダムに行うランダムフォレスト データの重みと決定木の信頼度を利用するブースティング
8	2022年2月9日(水) 19:00~20:30	機械学習のまとめと応用例	新田 克己 名誉教授 東京工業大学 情報理工学院 特任教授	<ul style="list-style-type: none"> 機械学習のまとめ 機械学習の補足：強化学習(環境からの報酬を最大化する行動の学習) 機械学習の応用例(次元圧縮の応用例) 潜在意味解析・トピックモデル

【演習】

No.	日程	演習課題名	講師	演習課題内容
1	課題提示 2022年2月9日(水) 課題提出期限 2022年2月28日(月)	演習課題1	柳澤 淡雨 助教 東京工業大学 情報理工学院 助教	<ul style="list-style-type: none"> 全8回の講義で得た知識を元に、実践的な演習を各自で取り組む 提出された課題を講師により添削・講評をし返却
2	課題添削返却 2022年3月31日(木)	演習課題2	安尾 信明 特任講師 東京工業大学 物質・情報卓越教育院 特任講師	