

お茶の水女子大学におけるデータサイエンス教育： 人文/社会科学・附属高等学校への展開

数理・データサイエンス教育強化拠点コンソーシアム
関東・首都圏ブロック 第5回ワークショップ

2020年12月23日





お茶の水女子大学は2019年

「文理融合AI・データサイエンスセンター」

を設立しました

文理融合AI・データサイエンスセンター



- AIおよびデータサイエンスに関係ある教育と研究を全学的に推進するための組織横断型のセンター

教育部門

文理融合データサイエンス
科目の全学的展開

データサイエンス科目群の
学際カリキュラム策定

データサイエンス教材の
協定校・附属校への提供

研究部門

学内共同研究や
産学連携の推進

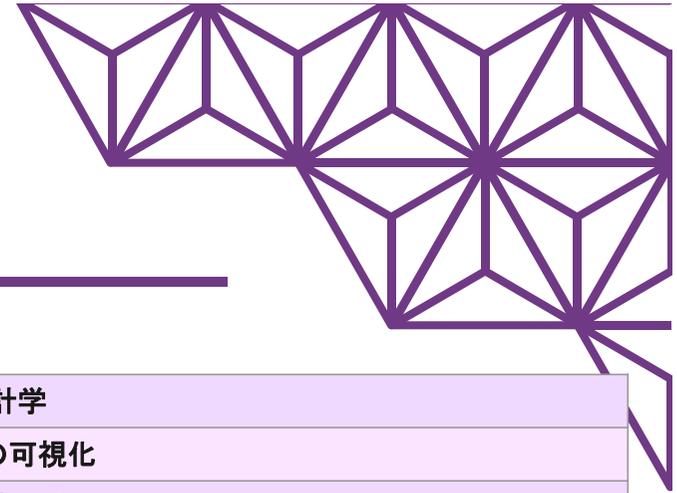
AI関連研究に従事
する学生の支援

データ管理環境・
知的財産の強化



「文理融合データサイエンスI」

コア科目(1年後期)... 文化現象を対象に



本授業の主題はデータサイエンスの手法を学ぶことに加え、文化現象を対象にデータサイエンスの手法を用いた事例を概観し、文理融合型の研究におけるデータ分析の有効性と重要性を学ぶことです。

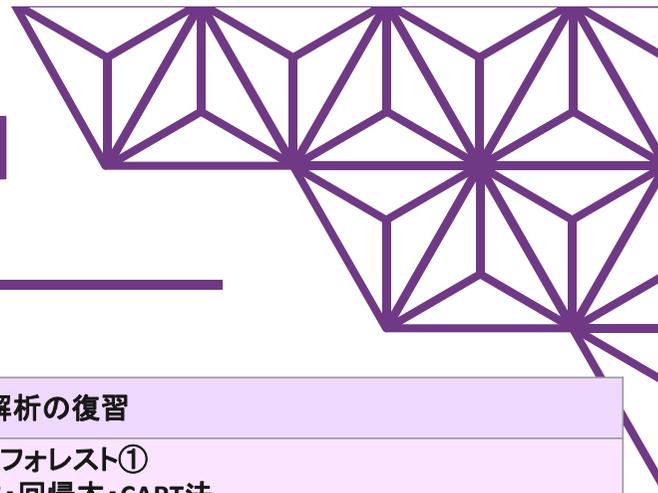
それ故、本授業ではデータサイエンスの手法を理解することだけでなく、文化に関するデータに対しデータサイエンスの分析手法をどのように応用するのか理解することが重要になります。

第1回	記述統計学
第2回	データの可視化
第3回	確率の考え方
第4回	推測統計学
第5回	さまざまな検定
第6回	事例(1): 区間推定による薬師寺の建立年代の推定
第7回	回帰分析1 事例(2): シミュレーションデータの線形回帰分析
第8回	判別分析 事例(3): 日本近代文学における自然主義作家の識別
第9回	回帰分析2 事例(4): ロジスティック回帰分析
第10回	主成分分析 事例(5): 主成分分析を用いた夏目漱石の小説の分析
第11回	対応分析 事例(6): クロス集計と対応分析
第12回	クラスター分析 事例(7): 小説のクラスタリングによる著者識別
第13回	樹木モデル 事例(8): 決定木を用いたシミュレーションデータ分析 事例(9): 決定木を用いた短い文章の書き手の判別
第14回	樹木モデルの応用 事例(10): 『源氏物語』と『うつほ物語』の特徴語の抽出
第15回	多変量解析の復習



「文理融合データサイエンスII」

コア科目(2年前期)... 文化現象を対象に



本授業の主題は文理融合データサイエンス I において学んだ基礎的な多変量解析の手法を復習し、新たに機械学習の手法を学ぶことです。

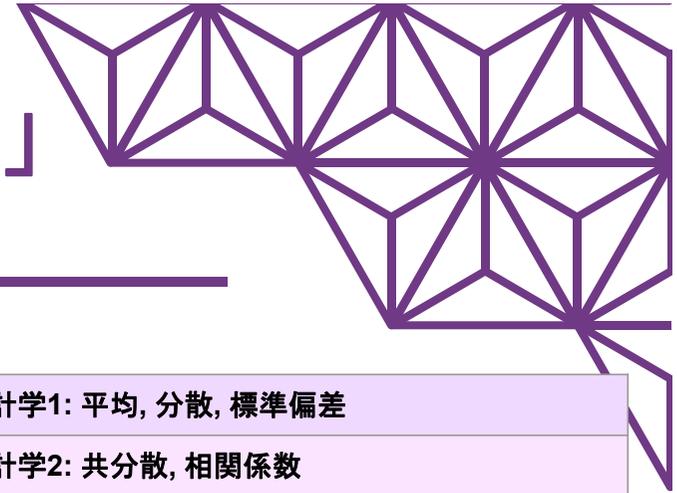
これに加えて、文化現象を対象とするデータサイエンスの手法を用いた研究事例を概観することで、文理融合型の研究におけるデータサイエンスの有効性と重要性を理解することが目標となります。

第1回	多変量解析の復習
第2回	ランダムフォレスト① 決定木・回帰木・CART法
第3回	ランダムフォレスト② ブートストラップ・分類問題と回帰問題・変数重要度
第4回	ランダムフォレストの演習
第5回	ネットワーク分析 有向グラフと無向グラフ・隣接行列・中心性
第6回	ネットワーク分析の演習
第7回	これまでの復習および中間レポートの説明
第8回	ニューラルネットワーク
第9回	ニューラルネットワークの演習
第10回	SVM ソフトマージンとハードマージン・カーネル法
第11回	SVMの演習
第12回	相関ルール
第13回	相関ルールの演習
第14回	深層学習の紹介
第15回	これまでの総復習



「文理融合データサイエンスIII」

コア科目(2年後期)...計量経済学への発展



本授業では統計学や計量経済学の理論的な知識を取得することと、ExcelやStataなどの計量ソフトを使用してデータ分析ができることを目的とします。

e-Statから政府統計を自分で加工してデータ分析を行うような課題を出します。

第1回	記述統計学1: 平均, 分散, 標準偏差
第2回	記述統計学2: 共分散, 相関係数
第3回	確率論の復習1: 条件付き期待値
第4回	確率論の復習2: 条件付き分散
第5回	確率論の復習3: 累積密度関数
第6回	推測統計学1: 正規分布, t分布
第7回	推測統計学2: 標本平均の平均と分散, t検定
第8回	回帰分析1: 単回帰分析, 最小二乗法
第9回	回帰分析2: 重回帰分析
第10回	回帰分析3: t検定とF検定
第11回	回帰分析4: Excelを使った演習
第12回	パネルデータ分析の基礎1: パネルデータの特徴
第13回	パネルデータ分析の基礎2: 固定効果法と変量効果法
第14回	パネルデータ分析の基礎3: パネルデータを用いた演習
第15回	まとめ

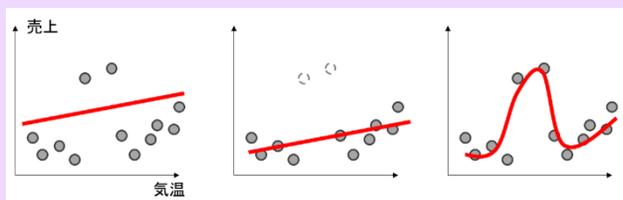
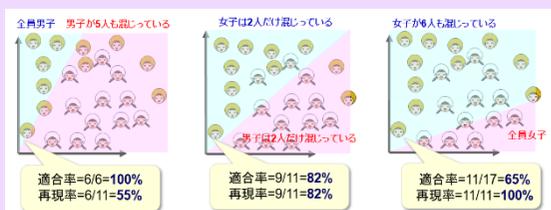


附属高校でのデータサイエンス教育

- SSH (Super Science High school) 1年生科目
「課題研究基礎」での講演と実習

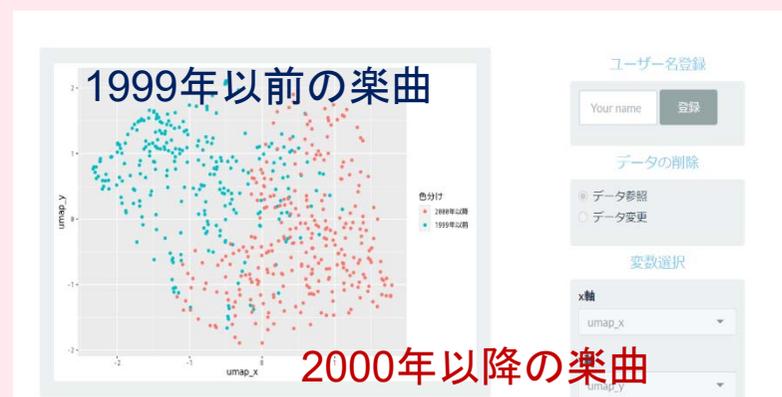
データサイエンスの講演

判別・回帰・可視化の概念を
実例や例題を交えて解説



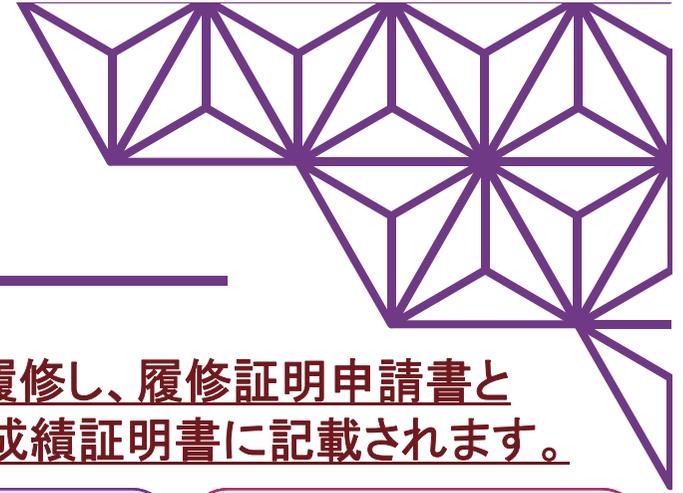
データサイエンスの実習

音楽の判別分析をウェブ上の
アプリケーションで体験実習



※2021年度以降では2年生「課題研究I」3年生「課題研究II」も支援予定

全学データサイエンス 学際カリキュラム



必修2科目＋基幹・連携8科目以上＝計10科目以上を履修し、履修証明申請書と
修了レポートを提出した学生には、履修証明書を授与し、成績証明書に記載されます。

人文・社会科学系

人文地理学分析基礎演習
地理情報システム演習
人間科学論 現代社会論
社会調査の設計と実施
学校社会学特殊講義
教育社会学特殊講義
社会調査法 応用統計学
応用生活統計学
社会統計学 心理統計法
生活調査法
生活社会調査実習

生活工学系

情報工学演習
計測工学 人間工学
電子工学 応用統計学
環境評価学
栄養疫学・統計

生命・物質科学系

物理実験学
計算物理学講義・演習
実験値解析法
計算化学 生物統計学
生命情報プログラミング演習

数学・情報系

グラフ理論 数理統計学
確率序論 データ解析序論
データ構造とアルゴリズム
コンピュータシステム序論
マルチメディア
情報理論 環境情報論
情報と職業 情報倫理

必修科目群:

文理融合データ
サイエンスI・II

基幹科目群:

基礎線形代数学／基礎微分積分学／統計学／
情報科学／情報処理学／メディアリテラシ／プログラミング演習1,2／
情報学演習1,2／文理融合データサイエンスIII・IV／デジタル概論

本日の講演内容



- **土山玄**
(お茶の水女子大学文理融合AI・データサイエンスセンター)
文学作品のテキストデータを題材としたデータサイエンス演習
- **松浦司**
(中央大学経済学部／お茶の水女子大学非常勤講師)
政府統計を用いたデータサイエンス演習の現状
- **伊藤貴之**
(お茶の水女子大学文理融合AI・データサイエンスセンター／理学部情報科学科)
附属高等学校でのデータサイエンス実習事例

