

知能システム開発特論 シラバス

授業情報

- 授業時間: 1 コマ 90 分
- 開講年次: 2017 年 秋
- 授業時間: 毎週火曜日 8:50~10:20
- 授業場所: ソフトウェア情報学部 A 棟 201 講義室
- 授業用 Web ページ: <http://p-www.iwate-pu.ac.jp/~david/ids>

授業のねらい

データサイエンスとは、主に大量のデータから汎用可能な知識を獲得する研究分野です。データサイエンティストであるためには、数学や統計学、機械学習、そしてデータベースなどコンピュータサイエンスにおける様々なスキルと、問題解決手法の理解を必要とします。この授業では、急速に発展するデータサイエンスに関連する研究を紹介しながら、それらの基本理論や手法、一般的な考え方の一部を身につけます。

学生はデータ収集・整形や探索的データ解析、予測モデリング、記述的モデリング、データ製品の生成、評価、そして効果的なコミュニケーションを含む実践的なデータサイエンスにおける様々な場面に対応するために必要な概念や技術、及び手法を学びます。

学修目標

7 回の講義と演習を通して、次のことができるようになることを想定している。

- データサイエンスとは何であるか説明する
- 統計的推論が意味することをかんたんに説明する
- 確率分布を特定し、モデルをデータに対し適合する
- R 言語を用いて基本的な統計的モデリングと解析を行う
- データサイエンスにおける探索的データ分析(EDA)の意義を説明する
- EDA とデータサイエンス処理を事例研究に適用する
- 基本的な機械学習アルゴリズムを適用する
- 特徴生成について区別する
- データマイニングのための基本的な数学的かつアルゴリズム的な手法を区別し説明する

成績評価の方法

- 出席(30%)
- 課題(30%)
- プレゼンテーション(40%)

テーマと授業概要

1. はじめに: データサイエンスとは何か?
2. 統計的推論
3. 探索的データ分析とデータサイエンスプロセス
4. 基本的な機械学習アルゴリズム
5. 特徴生成と特徴選択
6. データ可視化
7. テキストマイニング

学修にあたっての環境準備

RStudio、R プログラミング、もしくは Anaconda Python プログラミング

参考書等

- Data Science from Scratch (Joel Grus)
- ゼロからはじめるデータサイエンス - Python で学ぶ基本と実践 (Joel Grus (著)、菊池 彰 (翻訳))
- R によるデータサイエンス (金 明哲)