

| | | | | | |
|-----------|--|---------|---------|------|----|
| 科目名(英語名) | データサイエンス入門 (Introduction to Data Science) | | | | |
| 単位数 | 2 | 配当年次 | 1 | 開講時期 | 後期 |
| 対象コース | 全コース 学部専門選択 | | | | |
| 科目群 | | | | | |
| 全教員名 | 香取勇一、佐藤直行、他 | | | | |
| 授業・試験の形式 | | | | | |
| 授業概要 | データサイエンスは、膨大な量のデータから、新たな意味や価値、法則、関係性などを見つけ出す学問分野であり、自然科学、経済、医療などさまざまな分野の基盤となっている。この講義の前半では、統計解析向きプログラミング言語 (Python言語) を用い、課題・演習を通じてデータサイエンスにおける基本的な考え方を身につける。後半では、学術的もしくは実践的な観点において、より発展的なトピックスについて理解する。 | | | | |
| キーワード | コンピュータ科学, 情報科学, 統計学, プログラミング, データ分析 | | | | |
| 到達目標 | <ul style="list-style-type: none"> ・データサイエンスの基本的な考え方を理解する。 ・統計解析向きプログラミング言語 (Python言語) を用いて、データ解析を行うことができる。 ・データサイエンスの発展的なトピックスを理解する。 | | | | |
| 授業計画 | <p>第1～10回では、データサイエンスの基本的な考え方を身につける。</p> <ol style="list-style-type: none"> 1. データサイエンス概説, Pythonの基本操作 2. Pythonの基本 (条件分岐, 繰り返しなど) 3. 配列の基本操作 (numpy, ファイル入出力など) 4. データの可視化 (matplotlib, pandas) 5. 平均と分散, はずれ値 6. 相関と散布図, 相関係数 7. 確率分布と統計的推定 8. 教師あり学習・回帰分析 (線形回帰, 平均二乗誤差, 過学習など) 9. 教師あり学習・クラス分類 (ロジスティック回帰, 決定木など) 10. 教師なし学習 (クラスタリング, k-means法, 主成分分析など) <p>第11～15回は、各コースのデータサイエンス関連トピックスを紹介する。</p> <ol style="list-style-type: none"> 11. 例) 機械学習 (複雑系コース) 12. 例) 人工知能 (知能コース) 13. 例) 情報システムとデータ分析 (情報システムコース) 14. 例) 実験・社会調査データの分析 (デザインコース) 15. 例) 実社会におけるデータ分析 (ICTコース) | | | | |
| 事前・事後の学習 | <p>事前：授業内で指示された内容を予習してくる。</p> <p>事後：授業中に出席するレポート課題を行い、指示された内容を復習してくる。</p> | | | | |
| 成績の評価方法 | 講義内に行う小テスト, 課題 (レポート) により総合的に判断する。これらの比率は授業内に通知する。 | | | | |
| 教科書・参考書 | <p>教科書：「東京大学のデータサイエンティスト育成講座」 (マイナビ出版)。</p> <p>参考書：必要に応じて適時紹介する。</p> | | | | |
| 履修上の注意 | 4年次以上の履修は不可とする。 | | | | |
| DSOP対象コース | 全コース | DSOP科目群 | DS入門科目群 | | |
| 備考 | | | | | |