

## 入門科目 2単位取得

## ●データサイエンス概論 (人材像とそのレベルに達するためのステップ、そして基礎的概念を概説)(必修)

入門科目として、「データサイエンス概論」(必修)があります。この講義の目的は、教育目的とカリキュラムの全体像を学生に伝えることです。データサイエンス研究科修士課程において基礎となるデータエンジニアリング及びデータアナリシスに関する科目、さらにデータの特徴を表し分析の目的に適したモデルを構築するためのモデリング科目について概説します。さらに企業等の現場においてデータサイエンスを活かすために必要とされるプロジェクト

トマネジメントの方法論について講義します。また個人情報などのデータを扱う際の情報倫理についても補足します。これにより、社会や企業活動におけるデータサイエンスの重要性について理解し、データサイエンスの体系の概要と、データサイエンスを習得するためにどのような学習が必要とされるかについて理解することができます。

授業計画

- |                           |                               |
|---------------------------|-------------------------------|
| 1 教育目的とカリキュラムの関係          | 8 企業内データサイエンティストのリテラシー        |
| 2 上級データエンジニアリング           | 9 ソリューションのマネジメント              |
| 3 上級データアナリシス              | 10 ステークホルダーのマネジメント            |
| 4 モデリング方法論                | 11 データサイエンティストチームのマネジメントと人材育成 |
| 5 開発管理、プロジェクト管理           | 12 経営とデータサイエンス                |
| 6 グローバル企業におけるイノベーション管理    | 13~15 情報倫理                    |
| 7 企業におけるデータサイエンスの機会とその類型化 |                               |



## データエンジニアリング科目 2単位以上取得

- Webマイニング特論 (選択必修) / Webマイニング実践論
- サイバーフィジカル特論 (選択必修) / サイバーフィジカル実践論
- マルチメディア特論 (選択必修) / マルチメディア実践論

データサイエンスの一連の過程において、情報処理の専門知識とスキルは、3つのステージに分けられます。第1ステージはデータの取得と保存、第2は前処理および計算、第3はデータの可視化です。本研究科のカリキュラムでは、これらのステージで必要とされる専門知識を別々に学ぶのではなく、データの種類ごとに、第1ステージから第3ステージまでの全てを1つの科目の中で学びます。

具体的には、テキストデータを題材に全ステージを1人で担えるスキルを身につけるための講義として「Webマイニング特論」、さらに講義で学んだ内容を学生が自ら実践する力と自立的に学習し続けていく力を養うための実践形式の授業として「Webマイニ

ング実践論」を開講します。そして、センサーデータについては、「サイバーフィジカル特論」および「サイバーフィジカル実践論」、画像・音声データについては、「マルチメディア特論」および「マルチメディア実践論」において学びます。

一般にWebデータは大規模であるため、大規模計算を含む第2ステージの比重が大きくなります。そして、センサーデータについては、例えばスマートフォンやラズベリーパイなどを用いて自らデータを収集するため、第1ステージの比重が大きくなります。また、画像・音声データについては、可視化が重視されるため第3ステージの比重が大きくなります。

## データアナリシス科目 2単位以上取得

- モデリング基礎理論 (必修) / モデリング基礎実践論
- モデル評価論 / モデル評価実践論
- 確率過程理論 / 確率過程実践論

データを正しく分析するために、データに内在するランダムネスを処理・測定する専門知識とスキルを学びます。データアナリシス科目では、モデルを記述するための数学的道具立てとしてグラフィカルモデルや近似推論、そして欠損値および外れ値処理の理論的フレームワークを学ぶ「モデリング基礎理論」、情報量基準などの構築したモデルの良さを比較し評価する方法について学ぶ「モデル評価論」、金融データ等の時系列モデルに特化した数理を学ぶ「確率過程理論」の3つの講義があります。例えば、「確率過程理論」では、時々刻々と連続的に変化する不確実な現象を記述する数学モデルとして利用される確率過程について講義します。測度論からはじめ、それに基づいて確率論の基礎的な概念の定義や

諸性質を与え、極限定理やマルチンゲール理論など確率解析の基礎的事項を学びます。

これら3つの講義のいずれも学部レベルでは扱わない高度な統計理論やデータ分析法に関する講義によって構成されます。さらに、それぞれの講義について、学生が最新論文を輪講形式で紹介したり、すでに公表されているデータ解析結果を再分析し受講者で議論したりするような実践形式の授業「モデリング基礎実践論」、「モデル評価実践論」、「確率過程実践論」を用意します。これらはいずれも、モデリング科目において学ぶ各種モデリングの方法論を理解する上での基盤となります。

## モデリング科目 4単位以上取得

- 教師あり学習 (必修) / 教師あり学習実践論
- 教師なし学習 (必修) / 教師なし学習実践論
- 時系列モデリング / 時系列モデリング実践論
- 統計的モデリング / 統計的モデリング実践論
- 強化学習・転移学習 / 強化学習・転移学習実践論

データから価値創造をするためには、データの特徴を表し分析の目的に適したモデルを構築する必要があります。モデリング科目では、モデル化の方法論について、手法の種類ごとに最新の理論と方法を網羅的に学びます。具体的には、「教師あり学習」、「教師なし学習」、「時系列モデリング」、「統計的モデリング」、「強化学習・転移学習」の講義5科目と、実践形式で学生が最新の論文を紹介したり実装して実際のデータで評価したりすることにより自ら学ぶ力を養う「教師あり学習実践論」、「教師なし学習実践論」、「時系列モデリング実践論」、「統計的モデリング実践論」、「強化学習・転移学習実践論」の5科目を開講します。例えば、「教師あり学

習」では、現在大きな注目を集めている深層学習やスパースモデリング、またガウス過程、ベイズ最適化、隠れマルコフモデル等を学びます。「教師なし学習」では、ものづくりの現場で必要とされる異常検知技術、また行列分解、クラスタリングについても学びます。Pythonなどのプログラミング言語による演習も行い理解を深めます。「時系列モデリング」では、ARIMAモデルや状態空間モデル、モンテカルロフィルタなどを学びます。「統計的モデリング」では、混合モデルや一般化線形モデル、階層ベイズモデルについて学びます。

## 価値創造科目 10単位以上取得

- 意思決定とデータサイエンス (必修)
- 領域モデル実践論
- 課題研究 1、2、3、4 (必修)

データサイエンスの目的である価値創造のためには、統計学と情報学の専門知識だけでは不十分です。そこで、様々な領域の知識及び実践経験を身につけます。

まずは、データから価値創造するための「型」を身につける科目として、「意思決定とデータサイエンス」を開講します。この講義を通して、分析力から価値を生むには、分析問題を解く力だけでなく、それを意思決定の改善につなげる力も必要という意識改革を促します。実際のビジネスシーンを模した演習を通して、データ分析を意思決定に役立てる実践的な知恵を身につけます。データサイエンティスト協会の設定する「独り立ちレベル」から「棟梁レベル」の入り口に達するための橋渡しの役割を果たす科目です。

さらに多くの分野の領域知識に触れることのできる環境を用意するため、「領域モデル実践論」を開講します。領域知識を考慮せずに分析モデルを立てても有用な情報をデータから引き出すことは困難です。領域知識を活かして、モデルを構築する技量を向上させることができます。

そして、具体的な価値創造につなげる実践的研究を行うのが、「課題研究1」、「課題研究2」、「課題研究3」、「課題研究4」です。各種専門領域の担当教員の指導の下、大学の領域科学研究者、企業や自治体等と連携して実施される価値創造プロジェクト研究の一員として、現場の具体的な課題を読み取り、実際のデータを使って解析し、その知見を活かして価値創造を図ります。

## ビッグデータ解析等に基づく修士論文

複数の教員がチームを組み「課題研究1」から「課題研究4」の中で修士論文に関する研究指導を行います。学生は、興味と適性に基づいて、研究テーマを1年次春学期が終了するまでに決めます。それに応じて、担当教員と副担当教員が選ばれます。大学の領域科学研究者、企業や自治体等と連携して実施される価

値創造プロジェクト研究に参加して、データから価値を創造することを目指します。そして、その成果を修士論文としてまとめます。

所定の単位を取得した者に対し、次の学位基準に基づいて修士(データサイエンス)の学位を与えます。

**修士(データサイエンス)の学位  
提出された修士学位論文が  
右記の条件を満たすこと**

- データサイエンスおよびその関連分野における新たな成果を含む。
- 記述の論理構成が緻密であり、学問体系における成果の位置づけが明確で、かつ当該研究課題の周辺領域の専門家にも成果の意義が明快に伝わる。