データサイエンス応用コース

シラバス

2020年10月開講
全6ヶ月（20週、60時間）
毎週木曜日（18：15～21：15）
前半90分：講義　後半90分：演習
オンライン授業

応用I（18項目、9週）：10月第１週～12月第1週
ガイダンス、始業式　9月第4週

●データサイエンス入門
●情報理論の基礎
●機械学習の基礎

応用II（20項目、10週）：1月第1週～3月第3週
研究発表（ポスター）、修了式、講評 3月第４週

●マルチメディア
●データ活用
●データベース
●ソリューション企画
●データエンジニアリング

募集定員

20名（1クラス）

※参加者多数の場合は土曜午前も開講致します。

応用Ⅰ

データサイエンス入門

1.ニューラルネットワークの構造と学習（担当：高野）
【Keyword】ニューラルネットワーク(NN)，リカレントニューラルネットワーク(RNN)，教師あり学習，誤差逆伝搬法，シグモイド関数
【Abstract】ニューラルネットワークの構成要素と構造，及びニューラルネットワークのパラメータの最適化法に関して基礎的な数式を中心に講義する．

2. データ生成過程のモデル化（担当：高野）
【Keyword】生成モデル，識別モデル，内部状態，力学系，フィルタリング
【Abstract】生成モデルと識別モデルの差異，および生成モデルの一例を取り上げて，学習原理・適用方法について講義する．

情報理論の基礎

3. 情報源符号化（担当：中村）
【Keyword】符号化，情報エントロピー，データ圧縮
【Abstract】情報理論・符号理論の基礎について講義する．

4. データ構造（リスト，配列，木構造）（担当：中村）
【Keyword】データ構造，アルゴリズム，PとNP
【Abstract】コンピュータで用いられるデータ構造と，それを用いたアルゴリズムについて講義する．

5. 標本化・量子化（A/D変換）（担当：下川）
【Keyword】アナログ信号，コンバーター，シリアルデータ，サンプリング，直線性
【Abstract】A/D変換回路の基本的な仕組みと種類、様々な応用範囲について講義する。

6. 形式言語，形式手法（担当：下川）
【Keyword】オートマトン，状態遷移モデル，字句規則，構文規則，意味規則
【Abstract】言語の論理的記述及び理解のための方法について講義する。

7. アルゴリズム，数値計算（担当：中澤）
【Keyword】Python，繰り返し処理，条件分岐，関数，四則演算
【Abstract】Google ColabのJupiterNotebook（Python）機能を用いて，アルゴリズム（繰り返し処理，条件分岐，関数）と数値計算（ベクトル，行列，四則演算，固有値問題）の講義・演習を行う．

8. 自然言語処理，分散・並列コンピューティング（担当：中澤）
【Keyword】正規表現，勾配法，機械学習，GPU，MNIST
【Abstract】自然言語処理については正規表現のプログラミングを重点的に行い，分散・並列コンピューティングについてはGPUを用いた機械学習のPythonコードについて解説する．

機械学習の基礎

9. 弱いAI・強いAI（担当：中村）
【Keyword】AIの定義，弱いAI・強いAI，AI研究の歴史，チューリングテスト
【Abstract】人工知能の定義と研究の歴史について講義する．

10. フレーム問題（担当：中村）
【Keyword】エージェント，フレーム問題，記号接地問題，命題論理と述語論理
【Abstract】人工知能研究の抱える根本問題について議論する．

11. 探索・推論（担当：中村）
【Keyword】状態空間，探索の手法，論理と推論，論理プログラミング
【Abstract】問題解決のためのさまざまな手法を概観する．

12.知識表現（担当：中村）
【Keyword】知識ベース，プロダクション・システム，意味ネットワーク，不完全な知識
【Abstract】知識の表現形式について講義する．

13. データの分類：サポートベクターマシン・ランダムフォレスト（担当：高野）
【Keyword】識別関数，サポートベクターマシン，決定木，ブースティング，ランダムフォレスト
【Abstract】識別モデル（識別関数）の例を挙げながら，各モデルの学習方法およびアルゴリズムについて講義する．

14. 変分ベイズ法（担当：高野）
【Keyword】EMアルゴリズム，因子分解，サンプリング，点推定，ベイズ推定
【Abstract】EMアルゴリズムの復習とその問題点を抑えたうえで統計モデリングにおける変分ベイズ法について講義する．

15. ニューラルネットワークⅠ：教師あり学習（担当：下川）
【Keyword】パーセプトロン，隠れ層，誤差逆伝播法，画像認識
【Abstract】ニューラルネットワークの基本的な構造と学習のしくみについて講義する。

16. ニューラルネットワークⅡ：教師なし学習（担当：下川）
【Keyword】クラスタリング，隠れ層，オートエンコーダー
【Abstract】教師信号を使わず与えられたデータの構造から学習する方法と応用について講義する。

17. ディープラーニングⅠ：畳み込みニューラルネットワーク（担当：下川）
【Keyword】畳み込み，プーリング，パディング，全結合
【Abstract】ディープラーニングの設計でよく用いられる典型的な手法について講義する。

18. ディープラーニングⅡ：再帰型ニューラルネットワーク（担当：下川）
【Keyword】時系列データ，BPTT，LSTM，言語モデル
【Abstract】再帰型ニューラルネットワークを用いた自然言語処理について講義する。

応用Ⅱ

マルチメディア

19. スパースモデリング（担当：高野）
【Keyword】スパース表現，行列分解，OMP，MOD，k-SVD
【Abstract】スパース表現のためす数学モデルと目的関数に応じたモデルパラメータの最適化アルゴリズムや多変量解析手法との関係を講義する．

20. テキスト処理（担当：高野）
【Keyword】Bag-of-words，n-gram，トピック推定，形態素解析，翻訳モデル，
【Abstract】自然言語を数値として表現する方法，文章・文書の特徴量表現と統計的モデル化，その応用例について講義する．

21. 音声処理（担当：中村）
【Keyword】音声データ，音響モデル，言語モデル
【Abstract】音声処理に必要な音声データの基本概念について講義する．

22. 画像処理（担当：中村）
【Keyword】画像のデータ表現，画像解析手法
【Abstract】画像処理に必要な画像データの基本概念について講義する．

データ活用

23. Python入門（担当：下川）
【Keyword】Notebook，配列，ライブラリ，インタプリタ，リスト内包表記
【Abstract】Python の特徴及び基本的な使い方に関して講義する。

24. Python入門（続）（担当：下川）
【Keyword】Notebook，配列，ライブラリ，インタプリタ，リスト内包表記
【Abstract】Python の特徴及び基本的な使い方に関して講義する。

25. R言語入門（担当：上阪）
【Keyword】R・R studioの導入, データ型, 四則演算
【Abstract】Rの基本的な使い方について講義する．

26. R言語入門（続）（担当：上阪）
【Keyword】R, R studio, パッケージ, データの可視化
【Abstract】実際にRを使用して、基本的な使い方について演習を行う．

27. 特徴抽出（クラスタリング・グルーピング）（担当：上阪）
【Keyword】R, テキストデータ, クラスター分析
【Abstract】クラスター分析等の分析手法、テキストデータに対象とした応用例について講義する．

28. 特徴抽出（データマイニング）（担当：上阪）
【Keyword】R, テキストデータ, 特徴抽出
【Abstract】Rを用いてテキストデータに対し実際に演習を行う．

29. ビジネス活用（意思決定）（担当：松原）
【Keyword】機械学習の解釈性，説明性
【Abstract】機械学習モデルを現場に導入するには，そのモデルや結果が説明可能であることが求められる。解釈性，説明性に関する技術を概説する.

30. ビジネス活用（施策実施）（担当：松原）
【Keyword】因果推論，反実仮想
【Abstract】データには何らかのバイアスが存在し，それが因果効果の理解を困難にする.これへの対策として反実仮想を考慮した機械学習手法を取り上げ，概説する.

データベース

31. トランザクション処理（担当：中村）
【Keyword】トランザクション，ACID特性，OLTP，OLAP
【Abstract】トランザクション処理の基本概念について講義する.

32. 関係データベース設計と操作言語（担当：中村）
【Keyword】関係データベース，関係代数，SQL，正規化
【Abstract】関係データベースの概念とSQL言語の基本について講義する.

ソリューション企画

33. 要求分析，IT化対象の決定（担当：松原）
【Keyword】ビジネスプロセスモデリング，プロセス指向
【Abstract】業務フローの問題点の把握・共有には，その見える化が必要である。ビジネスプロセスのモデリング手法について概説する.

34.既存資産の再利用決定，ITソリューション（担当：松原）
【Keyword】ビジネスプロセス分析，PDCA
【Abstract】ビジネスプロセスの分析手法，プロセス改善のPDCAについて概説する.

データエンジニアリング

35. 構造化データ・非構造化ツールの統計解析ツール（担当：下川）
【Keyword】構造化タグ，マークアップ，データ検証，検索エンジン
【Abstract】構造化データ/非構造化データの違いと取り扱いについて概説する。

36. 蓄積，加工（担当：下川）
【Keyword】データベース，SQL，テーブル
【Abstract】テーブル設計、作成、リレーションの方法について概説する。

37. データウェアハウス（担当：下川）
【Keyword】データベース，スナップショット，再構成，集計，データマート
【Abstract】データベースの実運用の場面でよく利用する機能について紹介する。

38. 非構造化データ・データベース（担当：下川）
【Keyword】音声データ，画像データ，SNS，センチメント
【Abstract】非構造化データを使った最近の応用について紹介する。

応用コース

目標レベル：当該教育訓練が対象とする技術や手法等を活用して、業務上の課題の発見と解決をリードするとともに、後進育成にも貢献できるレベル

到達目標：
〇データサイエンス全般の知識を有し、機械学習を活用してデータに基づく意思決定を実践できる人材を育成する
・データサイエンス、情報理論、機械学習の基礎を習得し、問題解決に活用できる
・マルチメディア全般に関わる知識を習得し、問題解決に活用できる
・Python、R言語の理解からビジネス活用まで、データ活用に関する知識を習得し、問題解決に活用できる
・データベースの基本概念からデータの蓄積、加工までデータエンジニアリングに関する知識を習得し、問題解決に活用できる
・ビジネスプロセスモデリングに関する知識を習得し、問題解決に活用できる

修得できるスキル：
①データサイエンスに関する知識・技術を用いた問題解決実践
　データ指向の問題の理解、課題定義、分析手法の選択を行い、問題解決に向けた提案を行える
②Python、R言語を用いたデータ分析
　統計的な正しい理解のもとで、データ分析、データ可視化を行える
③データベース技術を用いたデータ蓄積、加工
　問題解決に向けたデータモデリング、およびデータの蓄積、加工を行える

到達目標把握方法：各回演習および研究発表における発表内容の審査

修了認定基準：
各回演習および研究発表のプレゼン資料、プレゼン内容、質疑応答への対応などを総合的に審査し決定。
審査では、①プレゼンテーション内容が所定の項目を網羅しているか、②正しいプロセスや手順で分析が実施されているか、
③分析結果について、正しい解釈や論理的な考察が展開できているかを総合的に判断し、採点する。修了に関する合否は担当教員の合議で判定する。

受講の利便性・サポート体制：
講義はオンデマンド配信。平日は繰り返し視聴可能。土曜日に時間を決めて再配信。
演習、研究発表は全てオンライン。欠席時は録画映像を配信。講座終了時まで復習可能。
全5回でオンラインによるスクーリングを実施し、受講生からの質疑応答を受け付ける。後日、全体に向けて録画を配信。
e-Learningは進捗管理、メンタリングボックスでの受講者からの質問を対応。
週に1回オフィスアワーを設定し、講師が随時相談に対応。
メール等でも随時質問対応。