

# 医療データサイエンス入門

[演習] 第2～4学年 後期 自由選択 2単位

《履修上の留意事項》毎回、各自のPCを必携のこと。授業では、毎回、各自のPCを使って演習をしますので、忘れずにPCを持参してください。  
「医療データサイエンス入門」を履修していること。「医療データサイエンス入門」と「医療データサイエンス入門」を共に履修した学生には「数理データサイエンスAI教育プログラム 修了証」カードを交付します。

《担当者名》教授/二瓶 裕之 nihei@ 講師/西牧 可織 nishimaki@

## 【概要】

歯科医師が活躍する医療分野においても重要となる「数理・データサイエンス・AI」を使いこなせる技能を修得する。「医療データサイエンス入門」では、文部科学省「数理・データサイエンス・AI教育プログラム(MDASH)」認定制度で指定された応用基礎コア「A I・データサイエンス実践」を扱う。模型都市を教材として使いながら、データエンジニアリングにおけるデータ収集・加工、学習、評価といった一連の流れを修得する。また、模型都市を対象とした「データ・A I活用 企画・実施・評価」の実践の場を通じて、数理・データサイエンス・A Iの活用における一連のプロセスの理解を深め、人や社会にかかわる具体的な課題の解決に活用できる能力を修得する。

なお、「授業内容および学修課題」に付記されている(3-9)は、数理・データサイエンス・A I(応用基礎レベル)モデルカリキュラムの学修項目の番号である。

## 【学修目標】

歯科医師業務に必要とされる「数理・データサイエンス・A I」の知識と、それを活用する技能を身につけるために、MDASH応用

基礎コア「A I・データサイエンス実践」を理解する  
数千件～数万件のデータを加工処理するプログラムを作成できる  
機械学習、深層学習、強化学習の基本的な概念を説明できる  
AI技術(学習、認識、予測・判断、知識・言語、身体・運動)を活用し、課題解決につなげることができる  
複数のAI技術が組み合わされたAIサービス/システムの例を説明できる学生生活とデータサイエンスの関連性について説明できる

## 【学修内容】

| 回 | テーマ                          | 授業内容および学修課題  | 担当者            |
|---|------------------------------|--|----------------|
| 1 | ニューラルネットワーク(NN)の構成           | ニューロン、シナプス：numpy, math<br>パーセプトロン、重み係数・バイアス、活性化関数：if, def, return<br>データ表現、データ収集：matplotlib, seaborn<br><br>【ニューラルネットワーク演習(1～3回目)】<br>模型都市に付置した人型模型の計測データや模型都市の環境データと、自ら実装したNNモデルを使いながら予測・判断を行う。 | 二瓶 裕之<br>西牧 可織 |
| 2 | 機械学習・深層学習                    | ビッグデータの収集<br>医療職のなかで進む機械学習の応用と発展<br>医療職のなかで進む深層学習の応用と発展につ<br>医療職において管理されているデータの利活用<br><br>【ニューラルネットワーク演習(1～3回目)】   | 二瓶 裕之<br>西牧 可織 |
| 3 | 予測・判断                        | 最小二乗法、損失：append, scatter<br>One-hot データ、正規化：to_categorical,<br>StandardScaler<br>モデルの構築と可視化、過学習<br>予測値の可視化<br><br>【ニューラルネットワーク演習(1～3回目)】   | 二瓶 裕之<br>西牧 可織 |
| 4 | ディープニューラルネットワーク(DNN)の構成(3-9) | 深層学習ライブラリ：TensorFlow, keras<br>入力層・隠れ層・出力層：Sequential<br><br>【ディープニューラルネットワーク演習(4～6回目)】  | 二瓶 裕之<br>西牧 可織 |

| 回  | テーマ                               | 授業内容および学修課題   | 担当者            |
|----|-----------------------------------|---|----------------|
|    |                                   | 模型都市に付置した人型模型の計測データや模型都市の環境データと、KerasによるDNNモデルを使いながら予測・判断を行う。   |                |
| 5  | DNNモデルの構築(3-9)                    | 層の指定：Dense, Activation, ReLU<br>モデルの構成：model.summary<br>モデルの作成：model.fit<br><br>【ディープニューラルネットワーク演習(4～6回目)】  | 二瓶 裕之<br>西牧 可織 |
| 6  | 予測・判断(3-9)                        | モデルの評価：model.evaluate<br>モデルの可視化：model.layers<br>モデルによる予測：model.predict<br><br>【ディープニューラルネットワーク演習(4～6回目)】   | 二瓶 裕之<br>西牧 可織 |
| 7  | 畳み込みニューラルネットワーク(CNN)の構成と人体構造(3-9) | 人の視覚の経路、単純型細胞と複雑型細胞<br>畳み込みニューラルネットワーク<br>ベクトルと行列<br>畳み込み層、プーリング層<br><br>【畳み込みニューラルネットワーク演習(7～9回目)】<br>模型都市の人型模型や建築物の画像データと、KerasによるCNNモデルを使いながら画像認識技術のスキルを学ぶ。  | 二瓶 裕之<br>西牧 可織 |
| 8  | CNNモデルの構築(3-9)                    | ビッグデータの収集<br>モデルの構築と作成<br>モデルの可視化と予測・評価<br><br>【畳み込みニューラルネットワーク演習(7～9回目)】   | 二瓶 裕之<br>西牧 可織 |
| 9  | 画像認識アプリの構築・運用(3-9)                | AIの開発環境と実行環境：ngrok、Auth token<br>フロントエンドアプリケーション：streamlit<br>画像認識アプリの開発<br><br>【畳み込みニューラルネットワーク演習(7～9回目)】  | 二瓶 裕之<br>西牧 可織 |
| 10 | AI・データサイエンス実践(データエンジニアリング基礎)      | 認識技術の活用事例：cv2(OpenCV)<br>VR映像からの動画像・静止画像収集<br><br>【AI・データサイエンス実践(データエンジニアリング基礎)演習(10～12回目)】<br>画像・動画像処理ライブラリ(OpenCV)を使って、模型都市のVR映像から取得した動画像や静止画像に対して輪郭抽出、特徴抽出、色検出などを行いながら、データ収集・加工、学習、評価といった一連の流れを修得する。 | 二瓶 裕之<br>西牧 可織 |
| 11 | 輪郭検出                              | カスケード分類：CascadeClassifier、detectMultiScale<br>エッジ検出：Canny<br>直線検出：HoughLines<br>円検出：HoughCircles<br>動画像(MPEG4)からの検出<br><br>【AI・データサイエンス実践(データエンジニアリング基礎)演習(10～12回目)】                                    | 二瓶 裕之<br>西牧 可織 |
| 12 | 特徴抽出・色検出                          | 特徴抽出：AKAZE<br>特徴点マッチング：drawMatches<br>色検出：cvtColor<br>動画像(MPEG4)からの抽出<br><br>【AI・データサイエンス実践(データエンジニアリング基礎)演習(10～12回目)】  | 二瓶 裕之<br>西牧 可織 |

| 回  | テーマ                              | 授業内容および学修課題   | 担当者            |
|----|----------------------------------|---|----------------|
| 13 | AI・データサイエンス実践（データ・AI活用 企画・実施・評価） | <p>物体検出：YOLO<br/>VR映像からの動画像・静止画像収集<br/>アノテーション：LabelImg, YAML</p> <p>【AI・データサイエンス実践（データ・AI活用 企画・実施・評価）演習(13～15回目)】<br/>物体検出ライブラリ(YOLO)とアノテーションツール(LabelImg)を使って、模型都市のVR映像から取得した動画像や静止画像に対して物体検出を行いながら、数理・データサイエンス・AIの活用における一連のプロセスの理解を深める。模型都市から得られる実データを使用することで、データ活用に取り組むまでの前処理に係る労力を実感するなど「手触り感」も含めて学修する。また、4～5名程度の履修学生がグループになってアノテーションを行って、グループ毎の観点から物体検出モデルを作り、最後に、オンライン上で、各グループで制作した物体検出映像を発表し合う。</p> | 二瓶 裕之<br>西牧 可織 |
| 14 | AIシステムの開発、テスト、運用                 | <p>モデルの構築：yolov5/train.py<br/>静止画像からの物体検出：yolov5/detect.py<br/>動画像からの抽出：yolov5/detect.py</p> <p>【AI・データサイエンス実践（データ・AI活用 企画・実施・評価）演習(13～15回目)】</p>   | 二瓶 裕之<br>西牧 可織 |
| 15 | AIシステムの品質、信頼性                    | <p>品質の検証<br/>信頼性の検証<br/>オンライン発表</p> <p>【AI・データサイエンス実践（データ・AI活用 企画・実施・評価）演習(13～15回目)】</p>  | 二瓶 裕之<br>西牧 可織 |

#### 【授業実施形態】

面接授業

授業実施形態は、各学部（研究科）、学校の授業実施方針による

#### 【評価方法】

授業参加態度（PBLの回には討議の参加態度）30%と毎回(15回)の授業で作成する課題の評価を70%で評価する。

#### 【教科書】

本学DX推進サイト (<https://dx.hoku-iryu-u.ac.jp/>) に公開している資料を使用する

東京大学 数理・情報教育研究センター ([http://www.mi.u-tokyo.ac.jp/6university\\_consortium.html](http://www.mi.u-tokyo.ac.jp/6university_consortium.html)) に公開されているリテラシーレベル教材と応用基礎レベル教材も使用する。

#### 【備考】

1. 学習教材（授業資料）の配信、学習課題の提示
  - ・ Google オンラインアプリケーションを利用して学習課題を提示する。
2. 授業に関する学生相互の意見交換やグループ学習の実践
  - ・ 学生相互の意見交換を目的にGoogleオンラインアプリケーションを活用する
3. 授業時間中にその場で学生の理解度を把握する技術の活用
  - ・ Google Formを活用し、授業時間中にその場で学生の理解度を把握する

#### 【学修の準備】

- ・ 自分で所有しているパソコンを利用して事前学修の項目を確認したり、授業で作成した課題を再度作成するなどの事後学習（復習）を行うこと（80分）。
- ・ SGDやPBLでは、事前に具体的な調査資料を配布するので、個々で調べて授業に臨むこと(80分)。授業終了後は、グループ討議の結果を自分なりに振り返り、授業中に作成したプロダクトを再度作成するなどの事後学修（復習）を行うこと(80分)。

#### 【ディプロマ・ポリシー(学位授与方針)との関連】

DP1. 安全で質の高い歯科医療を提供するために必要な専門知識に基づく問題解決能力と患者ケアのための診療技能とからなる専門的実践能力、および医療・医学研究の発展のために必要な情報・科学技術の活用能力を身につけている。

( 専門知識に基づいた問題解決能力、患者ケアのための診療技能、情報・科学技術を生かす能力 )

DP3. より安全で質の高い歯科医療を実践し社会に適応する医学を創造していくために生涯にわたって自己および他の医療者との研鑽を継続しながら医療者教育と学術・研究活動にも関与できる能力を身につけている。

( 科学的探究、生涯に渡ってともに学ぶ姿勢 )