

シラバス

授業科目名	年度	学期	開講曜日・時限	学部・研究科など	担当教員	配当年度	単位数
AI・データサイエンスツールⅣ	2021	後期	他		井田 佳祐、酒折 文武	1年度配 当	2

履修条件・関連科目等

履修者数の上限は、100名程度となります。

授業で使用する言語

日本語

授業で使用する言語（その他の言語名）

授業の概要

AI・データサイエンス全学プログラムの応用基礎レベルに位置づけられ、スキルを身につけるツール科目の中でも、今後の応用力につながるツール科目である。汎用的プログラミング言語 Python とデータベース言語 SQL を用いて、ビッグデータを処理・加工したり、基本的なデータ分析を行うことができる基本的なスキルを学んでいく。

前半は Python 言語の基礎を学ぶ。人工知能 (AI) 開発やデータサイエンスにおけるプログラミング言語としてももっとも広く活用されているのが Python である。Python はウェブシステムやアプリ開発にも用いられている汎用的なプログラミング言語で、可読性にすぐれ、機械学習や深層学習を行う環境も整っているため、多くのデータサイエンティストに活用されている。この科目の前半ではプログラミングにあまり馴染みのない学生を対象として、Python 言語の基礎と、Python を用いて基本的なデータ分析などを行う方法について学ぶ。

後半はデータベース言語である SQL を扱う。大規模なデータを収集し、使いやすい形で整理したものをデータベースという。SQLはデータベースを操作するための言語のひとつであり、「クエリ」と呼ばれる命令に基づいて大規模なデータを効率的に操作して、条件に見合うデータを検索・抽出したり、データの編集を行うことができる。SQLにも様々な種類があるが、この科目では SQLite という軽量のシステムを用い、データベースの作成、管理、そして活用を学ぶ。Python との連携についても学ぶ。

本科目は【動画配信型オンライン授業】（オンデマンド）であり、動画によるレクチャーと実習部分からなる。

実習部分については、動画あるいは静的なウェブページの内容を参考にして、自身の PC で同じことを実行しながら進め、適宜演習問題にチャレンジしながら理解を深める。わからないところは manaba のスレッドなどを活用して講師やTAによるサポートが受けられるようにする。小課題は、課題がクリアできていない場合にはTAのサポートを得ながら再提出することが必要である。

科目目的

AI・データサイエンスにおいて、データ分析・機械学習や可視化をウェブアプリと連動させたり、音声認識・画像認識などのタスクを行ったたり、さらには深層学習や強化学習を含む人工知能を開発したりするためには、汎用的なプログラミング言語が不可欠である。また、格納されているデータベースの操作を行うにはデータベース言語が必要となる。本科目の目的は、汎用的なプログラミング言語の中でデータサイエンスにおける中心的な役割を演じているPythonと、データベース言語SQLの基礎を理解することである。

到達目標

本科目の到達目標は、Pythonを用いて数十行程度のプログラムを読んで理解・修正できること、そして簡単なプログラムを自分で作成できること、および、SQLiteを用いてデータベースからのデータ抽出や基本的な処理を行えることである。

授業計画と内容

【動画配信型オンライン授業】（オンデマンド）

■ 前半：Python

第1回 データサイエンスとPython

- データサイエンス界隈の関連技術の紹介
- Pythonとは、Pythonでできること
- Google アカウントの作成とColaboratoryのセットアップ
- ノートブックの使い方
- 他のPython実行環境

第2回 プログラミング入門

- 変数、演算、代入
- リストとタプル

第3回 関数の基礎

- 数学関数
- 文字列関数
- 小課題1：与えられたプログラムがどのようなことを実行しているのかを詳しく説明する

第4回 分岐とループ

- if文による条件分岐
- for文による繰り返し

第5回 NumPyによるベクトル・行列演算

多次元配列の必要性

配列と要素型

配列要素の操作と演算

○小課題2：与えられた簡単なタスクを実行するプログラムを作成する

第6回 Matplotlibによるグラフ作成

基本的なグラフ作成（線グラフ、棒グラフ、ヒストグラム、散布図）と加工

第7回 pandasによるデータ分析の基礎

SeriesとDataFrame

csvファイルの読み込みからのDataFrameの作成

データの操作

基本統計量の算出

○前半レポート：データを読み込んで簡単な分析とグラフ作成を行い、その結果をまとめる

■後半：SQL

第8回 データサイエンスとデータベース

リレーショナルデータベース（RDB）とは

さまざまなSQL

SQLiteのダウンロードと設定

SQLiteでできること、他のSQLとの差異

第9回 データベースの作成と接続、テーブル操作

既存データベース（DB）との接続、簡単な操作

新規DBの作成

テーブルの作成（CREATE、DROP）

データの追加・更新・削除（UPDATE、DELETE、INSERT）

トランザクション制御

第10回 データの抽出

データの取得（SELECT）

条件によるデータの抽出（WHERE、BETWEEN、IN、LIKEなど）

ソート（ORDER BY）

○小課題3：指定された簡単なDBを作り、そこから情報を取り出すためのコードを作成する

第11回 テーブルの結合

複数のテーブルの結合（INNER JOIN、OUTER JOIN、CROSS JOIN）

第12回 DBからのデータ抽出と集計

グループごとの集計（GROUP BY）

これまでに学んだことを活用しての演習中心回

○小課題4：与えられた中規模のデータベースから情報を取り出し集計する

第13回 SQLite 関数の活用

count、sum、avg、date、length、replaceなどの紹介

第14回 PythonとSQLiteとの連携

sqlite3モジュールによるPython上でのSQLite操作（connect、execute）

○後半レポート：SQLiteデータベースからデータを抽出し集計、グラフ化し、結果を解釈する

授業時間外の学修の内容

授業終了後の課題提出

授業時間外の学修の内容（その他の内容等）

授業時間外の学修に必要な時間数／週

- ・毎週1回の授業が半期（前期または後期）または通年で完結するもの。1週間あたり4時間の学修を基本とします。
- ・毎週2回の授業が半期（前期または後期）で完結するもの。1週間あたり8時間の学修を基本とします。

成績評価の方法・基準

種別	割合 (%)	評価基準
レポート	80	<ul style="list-style-type: none">●前半（Python）レポート：40% 与えられたデータを読み込んで適切な分析と可視化を行い、結果を読み取ることができるかを問う。●後半（SQL）レポート：40% SQLite データベースから条件に見合うデータ抽出を行い、集計と可視化、結果の考察までを適切に行うことができるかを問う。
その他	20	小課題 20%（第3回、第5回、第10回、第12回の4回、各5点）。ただし、小課題は解けるまで再提出が必要であり、未提出の場合には単位認定しない。 小課題では、データ抽出、可視化、分析などの与えられたタスクを実行できるかどうかを問う。

成績評価の方法・基準（備考）

課題や試験のフィードバック方法

授業時間に限らず、manabaでフィードバックを行う

課題や試験のフィードバック方法（その他の内容等）

アクティブ・ラーニングの実施内容

実習、フィールドワーク

アクティブ・ラーニングの実施内容（その他の内容等）

授業におけるICTの活用方法

その他

授業におけるICTの活用方法（その他の内容等）

各自の PC で Python および SQL の実習を行う。

実務経験のある教員による授業

はい

【実務経験有の場合】実務経験の内容

（井田）株式会社ALBERTにてデータサイエンス業務およびデータサイエンス教育業務に従事。

【実務経験有の場合】実務経験に関連する授業内容

（井田）業務および教育業務で培った経験や技術に基づいて、データサイエンス分野で活用しうる実践力を養う授業を行う。

テキスト・参考文献等

テキストは指定しない。毎回資料を配布する。

その他特記事項

参考URL

コメント1

コメント2

コメント3

コメント4
