

学科	ITエンジニア科3年制		シラバス
コード	科目趣旨・目的(背景となる社会情勢・業界動向・資格試験との関係、および、カリキュラム全体における位置づけ)		
年度	2024	近年、AIが脚光を浴びており、AIを支える機械学習技術およびデータ分析技術が急速に発展・普及している。機械学習やデータ分析を活用する多くのプログラムはPython言語で記述されているため、産業界においてPython言語のスキルを有した技術者の確保が課題となっている。特に、Pythonでは機械学習やデータ分析のためのサードパーティ製ライブラリが充実しており、Pythonの標準ライブラリと同様に、サードパーティ製ライブラリを使いこなすことがPythonエンジニアの必須スキルとして求められている。以上の背景を踏まえて、本科目では、データ分析作業を支援するモジュールを提供しているライブラリであるpandasといったデータ分析の代表的なサードパーティライブラリの使い方を学び、様々なデータ分析プログラムを開発する力を鍛えることを目的とする。さらに、HTTPリクエスト・レスポンスを扱うためのライブラリであるrequestsやHTMLからデータを抽出できるライブラリであるBeautifulSoupを用いたWebスクレイピングについて学習し、Webスクレイピングからデータ分析までを行うプログラムを開発することで、実践的なデータ分析の知識とスキルを習得することを目的とする。	
学年	2		
期	後期		
分野名	基礎		
科目名	Python応用		
単位			
授業形態	科目概要		
実務連携型授業	○	本科目では、Pythonの代表的なサードパーティ製ライブラリであるNumPy, pandasを取り上げて、各種ライブラリの使い方を理解するとともに、データ分析のためのプログラムの開発に必要な知識およびスキルを習得する。さらに、Webの仕組みおよびWebスクレイピングについて学ぶことで、分析対象となるデータの収集に必要な知識およびスキルを習得する。Webスクレイピングでは、requestsやBeautifulSoupといったライブラリを用いて、Webサイトで公開されている情報の抽出を行う。また、抽出した情報をpandasを用いて分析することで、データの抽出から分析までの一連の流れを学習し、実践的なスキルの習得する。	
必修・選択	必修		
前提とする科目			
展開科目	キーワード		
関連資格	NumPy pandas データフレーム データクレンジング データ分析 Webスクレイピング HTMLフォーマット HTTPリクエスト・レスポンス requests BeautifulSoup		
教員		常勤	到達目標
サードパーティ製ライブラリであるNumPy, pandasを使いこなして、データ分析プログラムを記述できるようにすることを目標とする。また、Webスクレイピングに用いるライブラリであるrequestsやBeautifulSoupを活用してデータ収集プログラムを記述できるようにすることを目標とする。			
コマシラバス(90分授業コマ単位のシラバス)			
90分/コマ	コマ主題	コマシラバス項目	内容
1	データ処理	1.1 コマ主題細目	①CSVファイル ②CSVファイルの読み込み ③CSVファイルの書き込み ④CSVファイルの集計・可視化 ⑤NumPyによる配列の処理
		1.2 細目レベル	①CSVファイルについて解説する。 CSVは様々なアプリケーションで扱うことができる最も一般的なデータフォーマットの1つである。そのため、実務上CSV形式でデータを取り扱うことが多い。CSVの概要やデータフォーマットについて解説することで、CSVファイルを扱ったデータ処理の理解を促す。 ②PythonでのCSVファイルを読み込む方法について解説する。 csvモジュールのcsv.readerを使った方法について解説し、実装させる。 ③PythonでのCSVファイルにデータを書き込む方法について解説する。 csvモジュールのcsv.writerを使った方法について解説し、実装させる。 ④PythonでのCSVファイルを集計・可視化する方法について解説する。 CSVファイル内のデータを取り出し、集計を行い、そして集計した結果をコンソール上で可視化する方法について解説し、実装させる。 ⑤NumPyの配列とPythonのリストの相違点について説明した上で、NumPyの配列において、Pythonのリストと同様に使える機能を中心にして、サンプルプログラムを提示しながら解説する。また、NumPyの配列とPythonのリストを相互に変換する方法について、サンプルプログラムを提示しながら解説する。
		1.3 5キーワード	①CSVデータ ②csvモジュール ③csv.reader ④csv.writer ⑤区切り文字
		1.4 コマ要素	<input type="checkbox"/> 実務連携 <input type="checkbox"/> 理解度確認テスト <input checked="" type="checkbox"/> オリジナル教材 <input checked="" type="checkbox"/> ICT活用 <input checked="" type="checkbox"/> 実習・実技・実験・演習 <input type="checkbox"/> 該当なし

2	pandas(1)	2.1	コマ主題細目	①pandasの概要 ②pandasによるファイル読み込み ③pandasによるデータフレームの操作(基礎) ④pandasによる統計量の計算
		2.2	細目レベル	①データ分析作業を支援するモジュールを提供しているライブラリであるpandasの概要と、pandasが提供するデータフレームの概念について解説する。 ②pandasを用いてCSVおよびJSONファイルを読み込み、データフレームを生成する方法について、サンプルプログラムを提示しながら解説する。また、生成したデータフレームの内容を表示する方法について、サンプルプログラムを提示しながら解説する。 ③pandasを用いてデータフレーム内のデータを操作する方法について解説する。 具体的には、行や列を指定することによりデータを選択する方法、特定の条件を満たす行をフィルタリングする方法、特定の列を昇順や降順で並び替える方法、行や列をデータフレームに追加する方法、各行や各列に対して関数を適用することで値を変更する方法について、サンプルプログラムを提示しながら解説する。 ④データから算出できる平均値や中央値、最大値、最小値、標準偏差、標本サイズなどの各種統計量の意味を解説した上で、pandasを用いてデータフレームの各種統計量を計算する方法について、サンプルプログラムを提示しながら解説する。
		2.3	5キーワード	①pandas ②データフレーム ③データのフィルタリング ④データの並び替え ⑤統計量
		2.4	コマ要素	<input type="checkbox"/> 実務連携 <input type="checkbox"/> 理解度確認テスト <input checked="" type="checkbox"/> オリジナル教材 <input checked="" type="checkbox"/> ICT活用 <input checked="" type="checkbox"/> 実習・実技・実験・演習 <input type="checkbox"/> 該当なし
3	pandas(2)	3.1	コマ主題細目	①pandasによるデータフレームの操作(応用) ②pandasによる欠損値や異常値の確認 ③pandasによるデータクレンジング
		3.2	細目レベル	①pandasを用いてデータフレームを操作する応用的な手法について解説する。 具体的には、データフレーム内のデータを特定の条件でグループ化し集計する方法や、複数のデータフレームを結合して新たなデータフレームを作成する方法、データフレームの行と列を入れ替える方法などについて、サンプルプログラムを提示しながら解説する。 ②データの品質がデータ処理に与える影響の一例として、データに欠損値や異常値が含まれる場合に統計量やデータの傾向が大きく変化することを解説する。また、pandasを用いてデータフレーム内のデータの欠損値や異常値を確認する方法について、サンプルプログラムを提示しながら解説する。 ③データの品質を上げるためのデータクレンジングについて、欠損値や異常値の置き換えや削除、表記ゆれの補正などの手法を解説する。また、pandasを用いてデータクレンジングを行う方法について、サンプルプログラムを提示しながら解説する。 さらに、データクレンジングを行った後のデータの統計量が実際に変化していることについて、サンプルプログラムとその実行結果を提示しながら解説する。
		3.3	5キーワード	①pandas ②グループ化 ③欠損値 ④異常値 ⑤データクレンジング
		3.4	コマ要素	<input type="checkbox"/> 実務連携 <input type="checkbox"/> 理解度確認テスト <input checked="" type="checkbox"/> オリジナル教材 <input checked="" type="checkbox"/> ICT活用 <input checked="" type="checkbox"/> 実習・実技・実験・演習 <input type="checkbox"/> 該当なし
4	Webスクレイピング	4.1	コマ主題細目	①Webスクレイピング ②HTTPリクエスト・レスポンス ③HTMLのパーズおよびデータ抽出
		4.2	細目レベル	①分析対象のデータを収集する手法の一つとしてWebスクレイピングを紹介する。 Webスクレイピングとは、Webサイトで公開されている情報の中から、特定の情報を抽出する手法である。HTMLフォーマットを初めとするWeb技術を解説し、Webスクレイピングに必要な前提知識を修得させる。 ②HTTPリクエスト・レスポンスを解説したうえで、Pythonでリクエストを送信する方法、レスポンスを処理する方法について、HTTPリクエスト・レスポンスを扱うためのライブラリであるrequestsを用いて解説する。 そのうえで、教材で指定したWebサーバにリクエストを送信し、レスポンスとしてページのHTMLなどの情報を取得できることを確認させる。 ③Webスクレイピングの過程で取得できたHTMLをパーズし、目的のデータを抽出する方法について、PythonのWebスクレイピング向けライブラリであるBeautifulSoupを用いて解説する。 HTMLをパーズする際に用いるBeautifulSoupのメソッドについてもサンプルプログラムを提示しながら解説する。
		4.3	5キーワード	①Web技術 ②Webスクレイピング ③HTMLフォーマット ④HTTPリクエスト・レスポンス ⑤パーズ
		4.4	コマ要素	<input type="checkbox"/> 実務連携 <input type="checkbox"/> 理解度確認テスト <input checked="" type="checkbox"/> オリジナル教材 <input checked="" type="checkbox"/> ICT活用 <input checked="" type="checkbox"/> 実習・実技・実験・演習 <input type="checkbox"/> 該当なし

5	Webスクレイピングしたデータの加工・分析	5.1	コマ主題細目	①pandasを用いたWebスクレイピングデータの加工 ②pandasを用いたWebスクレイピングデータの分析
		5.2	細目レベル	<p>①前回講義でWebスクレイピングを用いて取得・抽出したデータをもとに、pandasライブラリを用いてデータフレームを作成し、分析が容易になるようにデータの加工を行うことで、後に分析を行うための前処理をする演習を実施する。</p> <p>本演習で、前回講義までに学習したWebスクレイピングおよびpandasライブラリによるデータフレーム操作の復習・応用の学習を行わせるとともに、Webスクレイピングしたデータを加工するまでの一連の流れをなぞることで、より実践的なスキルを習得させる。</p> <p>②①で加工したデータをもとに、pandasライブラリを用いて、欠損値の対処や表記ゆれの補正などのデータクレンジングを行ったうえで、データの集計や統計量の算出などの分析を行う演習を実施する。</p> <p>本演習で、前回講義までに学習したpandasライブラリの復習・応用の学習を行わせるとともに、実際にWebスクレイピングしたデータに対してデータクレンジングや集計・分析を行う一連の流れをなぞることで、より実践的なスキルを習得させる。</p>
		5.3	5キーワード	①Webスクレイピング ②pandas ③データ加工 ④データクレンジング ⑤データ分析
		5.4	コマ要素	<input type="checkbox"/> 実務連携 <input type="checkbox"/> 理解度確認テスト <input checked="" type="checkbox"/> オリジナル教材 <input checked="" type="checkbox"/> ICT活用 <input checked="" type="checkbox"/> 実習・実技・実験・演習 <input type="checkbox"/> 該当なし