

Grade	1	前期・後期	後期	単 位	1	科目分類	分子	
主担当教員	森河 良太			GPA	対 象		応用	選択総合（教職必須）
担当教員	野口 瑤						医科	
コンピテンシー	実行・課題解決・批判的思考・論理的思考							
履修前提								

ねらい

受講者は自分の所有するUnix系OSを搭載したノートパソコンを教材として毎回持参し、学内LANに接続できる講義室において実習的要素を取り入れた授業を受けます。そしてコンピュータが生命科学において果たす役割をデータサイエンスの知見に基づいて理解し、生物情報科学系の大規模なデータ分析に活用できる力を身に付けます。そのために例題実習を通してインタプリタ言語であるPythonを用いたプログラミングの基礎を学び、ターミナルによるコマンド操作に慣れ、計算科学のためのアルゴリズムの基礎を修得します。この授業で扱う内容を習得することで、現在急速に発達しているコンピュータサイエンスを生命科学の研究に役立てるための方法論と技術についての基礎的なスキルを身に付けます。

一般目標

生命科学の研究においてコンピュータの果たす役割とその基盤となる理論（アルゴリズム）を理解する。生命科学の研究におけるプログラミングの基礎概念を理解し、修得する。
実験データの統計解析やシミュレーションを行うための、データサイエンスに基づくコンピュータの活用方法を修得する。

授業内容

回数	担当	テーマ	到達目標
1	森河	ガイダンス：ターミナルの活用とviエディタ	Unix系CUIコマンドによる基本的なファイル操作ができる。viエディタの基本的な操作ができる。自分のパソコンのシェル環境を整え、Pythonを正しく起動できる。
2	森河	Python入門：対話モードとコーディング	機械語とプログラミング言語の関係、コンパイラとインタプリタの違いを説明できる。Pythonを対話モードで利用できる。エディタをPythonのソースコードを記述し、スクリプトとして実行できる。ファイルのアクセス権限について理解できる。
3	森河	Python入門：プログラムの基本構造	Pythonにおけるオブジェクトとクラスの概念を述べることができる。プログラミングにおける順接、反復、条件分岐について説明できる。論理和や論理積のような論理演算について述べることができる。処理の入れ子構造(ネスト)に応じた適切なプログラムを書ける。自然言語によって書かれたアルゴリズムを、Pythonを用いてプログラムに変換することができる。
4	森河	Python入門：ライブラリとデータの可視化	Pythonにおける関数の定義と外部ライブラリの利用方法について説明できる。数値計算ライブラリNumPyをインポートして四則演算を行うことができる。グラフ描画ライブラリmatplotlib.pyplotを使って折れ線グラフ、散布図、棒グラフを描ける。データ解析ライブラリpandasを使ってcsvファイルをプログラムに取り込み、データを集計できる。
5	森河	Python演習：配列の扱いとデータの入出力	Pythonにおけるリスト、タプル、辞書の概念について説明できる。NumPyにおける多次元配列データ構造ndarrayを説明し、使うことができる。数学使う関数のグラフをNumPyを使って描画できる。Pythonで計算したデータをテキストファイルに出力・入力することができる。
6	森河	Python演習：while文と方程式の数値的解法	elseやbreak, continueをwhile文の中で実装することができる。台形公式やシンプソン則、モンテカルロ法に基づいて関数の積分値を数値的に求める方法について説明でき、Pythonで実装できる。二分法とニュートン・ラフソン法を使って方程式の数値解を求めることができる。

回数	担当	テーマ	到達目標
7	森河	Python応用：差分方程式による微分方程式の解法	自然現象の定量的な解析において、微分方程式で記述することの重要性を認識する。生物の個体数の増加と指数関数、ロジスティック関数の関係について説明できる。人口増加の時系列データから、人間の知的活動を起点としたものの見方について認識する。1階常微分方程式の最も基本的な差分法であるオイラー法およびルンゲ・クッタ法について説明することができる。
8	森河	Python応用：if文とfor文を用いた配列の操作と行列計算	ndarrayを使ったベクトルと行列の演算子や関数について述べるができる。if文とfor文を使った行列式、小行列式、余因子、逆行列の計算アルゴリズムについて説明できる。再帰関数の意味と使い方について説明できる。連立方程式の計算方法について説明し、そのプログラムをPythonで作成できる。
9	森河	Python応用：実験データの処理と最小二乗法	実験やシミュレーションにおける数値誤差について説明できる。NumPyモジュールを使って1次データとしての実験データのばらつきや代表値、最小二乗法に関する統計処理ができる。実験データをPythonを使って誤差棒付きの散布図として表示できる。最小二乗法の原理を説明し、それを使って実験データを関数で表された曲線に近似できる。
10	野口	機械学習入門：教師あり学習と教師なし学習	人工知能の歴史、人工知能を利用した社会変革、人工知能の抱える問題について説明することができる。AIとビッグデータの関係、特化型AIと汎用型AIについて説明することができる。AIの最新技術の活用例（自然言語処置や画像認識）や仮説検証サイクルについて説明することができる。教師あり学習(回帰・分類)と教師なし学習(次元削減・クラスタリング)の違いについて説明することができる。モデル選択にまつわる精度と過学習、汎化性能について説明できる。
11	野口	教師なし学習：次元削減	教師なし学習の1つである次元削減について説明することができる。量的変数と質的変数の違いについて説明できる。データの集計と標準化、正規化について説明できる。散布図や散布図行列など特徴量間の関係を可視化する適切な軸の洗い出しができる。主成分分析を用いて、画像データを処理し特徴を取得することができる。Pythonのライブラリーの1つであるscikit-learnを用いて主成分分析を行うプログラムを作成し、実際にデータを解析することができる。多変量のビッグデータを可視化することができる。
12	野口	教師なし学習：クラスタリング	教師なし学習の1つである非階層型・階層型クラスタリングについて説明することができる。 dendrogramについて説明できる。Pythonのライブラリーの1つであるscikit-learnを用いてクラスタリングを行うプログラムを作成することができる。k-meansクラスタリングのアルゴリズムを理解し、プログラムを実装、実際にデータをクラスタリングすることができる。
13	野口	教師あり学習：分類	教師あり学習の1つである分類について説明することができる。学習データと検証データの必要性を説明できる。ダミー変数への変換について説明できる。決定木、ランダムフォレスト、SVMなど複数のモデルを作成できる。混同行列を用いて計算した4つの指標から、分類モデルの予測精度について説明することができる。ホールドアウト法や層化k分割交差検証を用いて汎用性の高いモデルを選択することができる。
14	野口	教師あり学習：線形回帰	Pythonでクラスを定義することができる。線形回帰である最小二乗法とRidge回帰、Lasso回帰の違いについて説明することができる。変数選択について説明することができる。Pythonのライブラリーの1つであるscikit-learnを用いて回帰を行うプログラムを作成することができる。Ridge回帰のアルゴリズムを理解し、プログラムを実装することができる。

準備学習：この授業は、各回において最低20分程度のオンラインによる予習が定められています。よって授業当日に配布するプリントは、前日までにCodexにてアップロードするので、よく読んでおいて下さい。また授業前に自分のノートパソコンを十分に整備しておいて下さい。また課題の他に、授業で学んだことを定着させるためにアンケート形式の「授業の振り返り」をCodexで提示するので、復習としてチャレンジして下さい。

授 業 形 式 : 自分のノートパソコンを用いた実習と講義を併用します。また授業中における課題の遂行では、ディスカッションを含めたグループワーク形式をとります。また、事前のビデオ配信による反転学習を行う回もあります。

課題に対する
フィードバックの方法
(試験やレポート等) : 課題レポートは主に大学のオンライン学習システム Codex にてフィードバックします。

評 価 方 法 : 期末試験は行わず、随時出される課題の提出やオンライン学習の成績(約70%)、授業中における課題の取り組み(約30%)等によって評価を行います。

教 科 書 : 井上英史 監修、森河良太、西田洋平、野口瑤 著、『基礎講義 情報科学 ~ デジタル時代の新リテラシーを身につける ~』、東京化学同人、2024年、ISBN:978-4-8079-2059-4。加えて Codex にて、授業内容に関する資料も配布します。

参 考 書 : 『新・明解Python入門』、柴田望洋著、SBクリエイティブ
『Pythonではじめる機械学習』、Andreas C. Muller著、オライリージャパン

その他特記事項 : 毎回、「ノートパソコン (MacBook)」と「電源アダプタ」を必ず持参して下さい。Windowsを搭載したノートパソコンの場合は、「情報科学I」でインストールしたWSL (Linux用Windowsサブシステム) の動作を確認しておいて下さい。また有線LAN に接続するための「LAN ケーブル」と「LAN アダプタ」は授業中に貸し出しますが、自分で持っている場合は持参して下さい。なお本授業は「情報科学」で学んだ知識や経験を前提として行います。
