

【リテラシーレベル】paizaラーニングの講座/問題集と数理・データサイエンス・AIモデルカリキュラム(2024年2月22日改訂)との対応表

http://www.mi.u-tokyo.ac.jp/consortium/model_literacy.html

数理・データサイエンス・AI教育強化拠点コンソーシアム

事務局：〒113-8656 東京都文京区本郷7-3-1 東京大学 数理・情報教育研究センター

| 1.社会におけるデータ・AI活用 | キーワード(知識・スキル) | paizaラーニング「講座」 | paizaラーニング「レベルアップ問題集」 |
|---------------------|--|---|-----------------------|
| 1-1. 社会で起きている変化 | <ul style="list-style-type: none"> ・ビッグデータ、IoT、AI、生成AI、ロボット ・データ量の増加、計算機の処理性能の向上、AIの非連続的進化 ・第4次産業革命、Society 5.0、データ駆動型社会 ・複数技術を組み合わせたAIサービス ・人間の知的活動とAIの関係性 ・データを起点としたものの見方、人間の知的活動を起点としたものの見方 | <p>プロンプトエンジニアリングの基礎 ChatGPT編1<生成AI></p> | |
| 1-2. 社会で活用されているデータ | <ul style="list-style-type: none"> ・調査データ、実験データ、人の行動ログデータ、機械の稼働ログデータなど ・1次データ、2次データ、データのメタ化 ・構造化データ、非構造化データ(文章、画像/動画、音声/音楽など) ・データ作成(ビッグデータとアノテーション) ・データのオープン化(オープンデータ) | | |
| 1-3. データ・AIの活用領域 | <ul style="list-style-type: none"> ・データ・AI活用領域の広がり(生産、消費、文化活動など) ・研究開発、調達、製造、物流、販売、マーケティング、サービスなど ・仮説検証、知識発見、原因究明、計画策定、判断支援、活動代替、新規生成など | <p>プロンプトエンジニアリングの基礎 ChatGPT編1<対話、コンテンツ生成、翻訳・要約・執筆支援、コーディング支援など生成AIの応用></p> <p>バイブコーディング入門 Claude Code編1: タッチボードアプリを作成しよう<対話、コンテンツ生成、翻訳・要約・執筆支援、コーディング支援など生成AIの応用></p> | |
| 1-4. データ・AI活用のための技術 | <ul style="list-style-type: none"> ・データ解析: 予測、グルーピング、パターン発見、最適化、モデル化とシミュレーション・データ同化など ・データ可視化: 複合グラフ、2軸グラフ、多次元の可視化、関係性の可視化、地図上の可視化、挙動・軌跡の可視化、リアルタイム可視化など ・非構造化データ処理: 言語処理、画像/動画処理、音声/音楽処理など ・特化型AIと汎用AI、今のAIで出来ることと出来ないこと、AIとビッグデータ ・認識技術、ルールベース、自動化技術 ・マルチモーダル(言語、画像、音声など)、生成AIの活用(プロンプトエンジニアリング) | | |
| 1-5. データ・AI活用の現場 | <ul style="list-style-type: none"> ・データサイエンスのサイクル(課題抽出と定式化、データの取得・管理・加工、探索的データ解析、データ解析と推論、結果の共有・伝達、課題解決に向けた提案) ・教育、芸術、流通、製造、金融、サービス、インフラ、公共、ヘルスケア等におけるデータ・AI活用事例紹介 | | |
| 1-6. データ・AI活用の最新動向 | <ul style="list-style-type: none"> ・AI最新技術の活用例(深層生成モデル、強化学習、転移学習、生成AIなど) ・AI等を活用した新しいビジネスモデル(シェアリングエコノミー、商品のレコメンデーションなど) ・基盤モデル、大規模言語モデル、拡散モデル | | |

| 2. データリテラシー | キーワード(知識・スキル) | paizaラーニング「講座」 | paizaラーニング「レベルアップ問題集」 |
|---------------|---|--|-----------------------|
| 2-1. データを読む | ・データの種類(量的変数、質的変数) | | |
| | ・データの分布(ヒストグラム)と代表値(平均値、中央値、最頻値) | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #04:画像から特徴量を抽出しよう<ヒストグラム> Pythonデータ分析入門編6: 記述統計量: #02:最大値、最小値、平均値、中央値<平均値、中央値> | |
| | ・代表値の性質の違い(実社会では平均値≠最頻値でないことが多い) | | |
| | ・データのばらつき(分散、標準偏差、偏差値)、外れ値 | Pythonデータ分析入門編6: 記述統計量: #03:分散、標準偏差<分散、標準偏差> | |
| | ・相関と因果(相関係数、擬似相関、交絡) | | |
| | ・観測データに含まれる誤差の扱い | | |
| | ・打ち切りや脱落を含むデータ、層別の必要なデータ ・母集団と標本抽出(国勢調査、アンケート調査、全数調査、単純無作為抽出、層別抽出、多段抽出) ・クロス集計表、分割表、相関係数行列、散布図行列 ・統計情報の正しい理解(誇張表現に惑わされない) | Pythonデータ分析入門編4: Jupyter Notebookとデータの読み書き #02:pandasによるデータの読み込み1<脱落を含むデータ> | |
| 2-2. データを説明する | ・データ表現(棒グラフ、折線グラフ、散布図、ヒートマップ、箱ひげ図) | Pythonデータ分析入門編5: #01データの可視化:pandasによるデータの可視化<棒グラフ、折線グラフ> Pythonデータ分析入門編5: #02:matplotlibによるデータの可視化1<棒グラフ、折線グラフ> Pythonデータ分析入門編5: #03:matplotlibによるデータの可視化2<棒グラフ、折線グラフ> Pythonデータ分析入門編5: #05:seabornによるデータの可視化<棒グラフ、折線グラフ、散布図> | |
| | ・データの比較(条件をそろえた比較、処理の前後での比較、A/Bテスト) ・不適切なグラフ表現(チャートジャンク、不必要な視覚的要素) ・優れた可視化事例の紹介(可視化することによって新たな気づきがあった事例など) ・相手に的確かつ正確に情報を伝える技術や考え方(スライド作成、プレゼンテーションなど) | | |
| 2-3. データを扱う | ・データの取得(機械判読可能なデータの作成・表記方法) ・データの集計(和、平均) ・データの並び替え、ランキング ・データ解析ツール(スプレッドシート、BIツール) | Pythonデータ分析入門編6: 記述統計量<和、平均> Pythonデータ分析入門編2: Seriesの基本<並び替え> Pythonデータ分析入門編4: Jupyter Notebookとデータの読み書き<データ解析ツール> | |
| | ・表形式のデータ(csv) | Pythonデータ分析入門編3: DataFrameの基本<表形式のデータ> Pythonデータ分析入門編4: Jupyter Notebookとデータの読み書き<csv> PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #05:scikit_learnで学習と予測を行う<csv> | |

| 3.データ・AI利活用における留意事項 | キーワード(知識・スキル) | paizaラーニング「講座」 | paizaラーニング「レベルアップ問題集」 |
|-----------------------|--|--|-----------------------|
| 3-1. データ・AIを扱う上での留意事項 | <ul style="list-style-type: none"> ・倫理的・法的・社会的課題(ELSI:Ethical, Legal and Social Issues) ・個人情報保護、EU一般データ保護規則(GDPR)、忘れられる権利、オプトアウト | | |
| | <ul style="list-style-type: none"> ・データ倫理:データのねつ造、改ざん、盗用、プライバシー保護 | 生成AI組み込みアプリの企画・開発 基礎編2: LLMを用いたプロダクト開発 #02: LLM導入の事業機会とリスク<プライバシー保護> | |
| | <ul style="list-style-type: none"> ・AI社会原則(公平性、説明責任、透明性、人間中心の判断) ・データバイアス、アルゴリズムバイアス ・AIサービスの責任論 | | |
| | <ul style="list-style-type: none"> ・データガバナンス | 生成AI組み込みアプリの企画・開発 基礎編2: LLMを用いたプロダクト開発 #13:セキュリティとガバナンス<データガバナンス> | |
| | <ul style="list-style-type: none"> ・データ・AI活用における負の事例紹介 | | |
| | <ul style="list-style-type: none"> ・生成AIの留意事項(ハルシネーションによる誤情報の生成、偽情報や有害コンテンツの生成・氾濫など) | 生成AI超入門 プロンプトエンジニアリングの基礎 ChatGPT編 #14:LLMの創作力とハルシネーション<ハルシネーション> | |
| 3-2. データを守る上での留意事項 | <ul style="list-style-type: none"> ・情報セキュリティの3要素(機密性、完全性、可用性) | テクノロジー編12: 情報セキュリティ(技術)<機密性、完全性、可用性> | |
| | <ul style="list-style-type: none"> ・匿名加工情報、暗号化と復号、ユーザ認証と、パスワード、アクセス制御、悪意ある情報搾取 | テクノロジー編12: 情報セキュリティ(技術)<暗号化と復号、ユーザ認証とパスワード、アクセス制御、悪意ある情報搾取> | |
| | <ul style="list-style-type: none"> ・情報漏洩等によるセキュリティ事故の事例紹介 | | |
| | <ul style="list-style-type: none"> ・サイバーセキュリティ | | |

| 4. オプション | キーワード(知識・スキル) | paizaラーニング「講座」 | paizaラーニング「レベルアップ問題集」 |
|----------------|---|--|---|
| 4-1. 統計および数理基礎 | <ul style="list-style-type: none"> ・確率、順列、組み合わせ | テクノロジー編02: コンピュータで役に立つ数学について 学んでいこう<確率、組み合わせ> | |
| | <ul style="list-style-type: none"> ・線形代数(ベクトル、ベクトルの基本的な演算、ノルム、行列とベクトルの積、行列の積、内積) ・1変数関数の微分と積分 | | |
| | <ul style="list-style-type: none"> ・集合、ベン図 | テクノロジー編02: コンピュータで役に立つ数学について 学んでいこう<集合、ベン図> | |
| | <ul style="list-style-type: none"> ・指数関数、対数関数 | | |
| 4-2. アルゴリズム基礎 | <ul style="list-style-type: none"> ・アルゴリズムの表現(フローチャート、アクティビティ図) | | |
| | <ul style="list-style-type: none"> ・並び替え(ソート) | 新・アルゴリズムとデータ構造入門 Python編4: 素朴なソートアルゴリズム<並び替え> 新・アルゴリズムとデータ構造入門 Python編5: 効率的なソートアルゴリズム<並び替え> Pythonデータ分析入門編3: DataFrameの基本: #07: ソート<並び替え> | 素朴なソートアルゴリズムメニュー<並び替え> 効率的なソートアルゴリズムメニュー<並び替え> ソートメニュー応用編<並び替え> |
| | <ul style="list-style-type: none"> ・探索(サーチ) | 新・アルゴリズムとデータ構造入門 Python編9: 二分探索<探索> | 線形探索メニュー<探索> 二分探索メニュー<探索> 幅優先探索・深さ優先探索メニュー<探索> 線形探索メニュー応用編<探索> 二分探索メニュー応用編<探索> 幅優先・深さ優先探索メニュー応用編<探索> 二分探索関連アルゴリズムメニュー<探索> |
| | | | |

| | | | |
|----------------------|--|---|---|
| 4-3. データ構造とプログラミング基礎 | ・数と表現、計算誤差、データ量の単位、文字コード、配列 | テクノロジー編01: コンピュータのデータ表現<数と表現、計算誤差、データ量の単位> | 配列メニュー<配列> 配列活用メニュー<配列> |
| | ・変数、代入、繰り返し、場合に応じた処理 | Python体験編1: Pythonをはじめよう: #07:変数にデータを入れる<変数、代入> | 算術・代入演算メニュー<代入> |
| | | Python体験編1: Pythonをはじめよう: #08:データを受け取る<変数、代入> | ループメニュー1<繰り返し> |
| | | Python体験編1: Pythonをはじめよう: #10:条件に一致したら処理を実行する場合に応じた処理 | ループメニュー2<繰り返し> |
| | | Python体験編1: Pythonをはじめよう: #11:条件に合わせて処理を変える場合に応じた処理 | 二重ループメニュー<繰り返し> |
| | | Python体験編1: Pythonをはじめよう: #12:数値を分類する場合に応じた処理 | 条件式メニュー<場合に応じた処理> |
| | | Python体験編1: Pythonをはじめよう: #15:複数データを分類する<繰り返し、場合に応じた処理> | 条件分岐メニュー<場合に応じた処理> |
| | | 新・Python入門編4: 変数を学習しよう<変数、代入> | プログラミング筋トレドリル4: 変数<変数> |
| | | 新・Python入門編10: 条件分岐を学習しよう<場合に応じた処理> | プログラミング筋トレドリル14: 定数回のループ<繰り返し> |
| | | 新・Python入門編11: ブール型を学習しよう<場合に応じた処理> | プログラミング筋トレドリル15: 様々なループ<繰り返し> |
| | | 新・Python入門編12: while文を学習しよう<繰り返し> | プログラミング筋トレドリル7: 条件分岐と等価演算子<場合に応じた処理> |
| | | 新・Python入門編13: for文を学習しよう<繰り返し> | プログラミング筋トレドリル9: 条件分岐と関係演算子<場合に応じた処理> |
| | | 4-4. 時系列データ解析 | ・時系列データ(トレンド、周期、ノイズ) ・季節調整、移動平均 |
| 4-5. 自然言語処理 | ・形態素解析、単語分割、ユーザ定義辞書、n-gram言語モデル、文章間類似度 ・かな漢字変換の概要 | | |
| 4-6. 画像認識 | ・画像データの処理 | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #04:画像から特徴量を抽出しよう<画像データの処理> | |
| | ・画像認識、画像分類、物体検出 | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #06:特徴量に明度のヒストグラムを利用しよう<画像データの処理> | |
| 4-7. データハンドリング | ・データベース(リレーショナルデータベース、SQL) | 新・SQL入門編<データベース> テクノロジー編10: データベースについて学習しよう(1)<データベース> テクノロジー編11: データベースについて学習しよう(2)<データベース> | 新・SQL入門編02ドリル<データベース> 新・SQL入門編03ドリル<データベース> 新・SQL入門編04ドリル<データベース> |
| | ・データクレンジング:外れ値、異常値、欠損値の処理 | Pythonデータ分析入門編9: データの集約:<欠損値> Pythonデータ分析入門編10: クリーニング:<欠損値> Pythonデータ分析入門編11: データ分析の実践:<外れ値、欠損値> | |
| | ・プログラミング(Python、R等) | 新・Python入門編<プログラミング> | |
| | ・データの抽出 | | |
| | ・データの結合 | Pythonデータ分析入門編7: データフレームのマージと連結:<内部結合、外部結合> | |
| | ・名寄せ | | |
| | ・生成AIを活用したデータ加工 | | |

| | | | |
|----------------------|--|--|--|
| 4-8. データ活用実践(教師あり学習) | ・教師あり学習による予測(例)売上予測、罹患予測、成約予測、離反予測など | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #03:問題と入出力データを考えよう<教師あり学習による予測> | |
| | | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #05:scikit learnで学習と予測を行おう<教師あり学習による予測> | |
| | | PythonxAI・機械学習入門編1: 機械学習の概要を知ろう #06:特徴量に明度のヒストグラムを利用しよう<教師あり学習による予測> | |
| | ・データの収集(分析に必要なデータの確認、対象となるデータの収集) | Pythonデータ分析入門編4: Jupyter Notebookとデータの読み書き<データの収集> | |
| | ・データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) | Python データ分析入門編1: データ分析を始めよう<データの加工> | |
| 4-9. データ活用実践(教師なし学習) | ・データの分析(単回帰分析、重回帰分析、ロジスティック回帰分析、モデルの評価) | | |
| | ・データ分析結果の共有、課題解決に向けた提案 | | |
| | ・教師なし学習によるグルーピング(例)顧客セグメンテーション、店舗クラスタリング | | |
| | ・データの収集(分析に必要なデータの確認、対象となるデータの収集) | | |
| | ・データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) | | |
| | ・データの分析(階層クラスタリング 非階層クラスタリング) | | |
| | ・データ分析結果の共有、課題解決に向けた提案 | | |