

分類番号

007

テーマ別調べ方ガイド



## Pythonによる 機械学習 について調べる



- ★ 関連キーワード
- ・プログラミング言語
  - ・データ処理
  - ・機械学習

### 「Paste (ぱすて)」とは？

ぱっと分かって、すっと頭に入る、テーマ別調べ方ガイドです。みんなの学習をサポートする、総合図書館ラーニング・サポーター（LS）による作成です。レポート作成の際などにお役立てください

## 1. イントロダクション

### 1-1. 「Python」と「機械学習」とは？

Python は、読みやすい文法と充実した標準ライブラリを備えたプログラミング言語です。数値計算やデータ分析、Web 開発、人工知能(AI)など幅広い分野で利用されており、学習しやすさと実用性の高さから初学者にも人気があります。

機械学習(Machine Learning)は、与えられたデータから規則や構造を自動的に抽出し、未知のデータに対して予測や意思決定を行うアルゴリズムの総称です。アルゴリズムは大きく三つに分けられます。第一に「教師あり学習」があり、これは過去データと正解(ラベル)を用いて数値を予測したりクラスを判定したりする手法です。第二に「教師なし学習」があり、これはラベルのないデータを入力としてクラスタリングや次元削減を行い、隠れた構造を把握するという手法です。第三に「強化学習」があり、エージェントが環境と相互作用しながら報酬を最大化する行動方策を学習します。近年はニューラルネットワークを多層化した「深層学習」により画像認識や自然言語処理が飛躍的に向上しました。

### 1-2. 学習するにあたってのポイント

- まずは基本文法を押さえる：変数、条件分岐(if)、繰り返し(for/while)、関数などを一通り理解すると学習が進めやすくなります。
- データ構造に慣れる：リスト、辞書、タプル、集合などを扱えるようになると、データ処理やアルゴリズム実装が安定します。
- ライブラリを使って実データを触る：NumPy・pandas での前処理、Matplotlib での可視化を通じて「データの形」を掴むのが効果的です。
- 小さな機械学習課題で一連の流れを回す：簡単な分類・回帰(例：花の分類、住宅価格予測など)を題材に、前処理→学習→評価→改善のサイクルを実装して経験すると定着が速くなります。

## 2. 基本要素

### 2-1. データ構造

- リスト型(list)：複数の要素を順番に格納する可変(変更可能)な配列。要素の追加・削除・並べ替えができる、データの収集や前処理で頻繁に使います。
- タプル型(tuple)：複数の要素を順番に格納する不变(変更不可)な配列。設定値や「変更されたくないデータ」を扱うのに向いています。
- 集合型(set)：重複を許さない要素の集まり。共通要素(積集合)や差分(差集合)などの集合演算ができる、データの重複排除や一致判定に便利です。
- 辞書型(dict)：キー(key)と値(value)の組でデータを管理する連想配列。属性情報や特徴量、設定パラメータなどを扱うのに適しています。

### 2-2. ライブラリ

- NumPy：多次元配列による高速な数値計算。

- pandas: DataFrame でデータの前処理(欠損・結合・集計・時系列)。
- scikit-learn: 機械学習一式(前処理／学習／評価)。
- OpenCV: 画像・動画処理(読み込み、変換、特徴抽出、前処理)。
- Matplotlib: 基本的な可視化(各種プロット)。
- PyTorch: 深層学習フレームワーク(GPU 対応)。
- TensorFlow / Keras: 深層学習(Keras は高水準 API)。

### 3. 学習用資料

#### 3-1. 最初に読むべき資料: 教科書

##### ■ [Python チュートリアル / Guido van Rossum 著 ; 鴨澤眞夫訳 \(第 4 版\)](#)

Python の作者による公式入門書であり、Python の基本文法やデータ構造、制御構文などを体系的に解説しています。初心者が Python の「文法の仕組み」と「設計思想」を理解するのに最適です。モジュールやクラスなどの基礎から、例外処理やテストコードまで幅広くカバーしており、Python 3.9 ドキュメントに準拠した改訂版です。

【書誌 ID= 2004511623】配架場所: 理工学図-東館 2F 図書 007.64||VAN

##### ■ [入門 Python 3 / Bill Lubanovic 著 ; 長尾高弘訳 \(第 2 版\)](#)

データサイエンス、ウェブ開発、セキュリティなど幅広い分野で活用される Python の入門書です。プログラミング初学者を対象に、Python の基本文法から応用的なプログラミング手法までを丁寧に解説しています。ウェブ開発、データベース操作、ネットワーク通信、並行処理などの実践的な内容も扱い、f 文字列など Python 3.9 の機能にも対応。Python の主要機能を網羅しており、学習用にもリファレンスとしても有用です。

【書誌 ID= 2004514548】配架場所: 総合図-A 棟 3 階 学習用図書 007.64||LUB

##### ■ [プログラミングの基礎からエラー処理、テストコードの書き方まで / Eric Matthes 著 ; 鈴木たかのり, 安田善一郎訳 \(改訂新版\)](#)

世界的ベストセラー『Python Crash Course』の日本語改訂版です。Python 3.11 対応のほか、開発環境を Visual Studio Code に更新。変数、リスト、if 文、クラスなどの基礎構文を短期間で習得できるよう構成されています。プログラミング初心者や、実践的なコードスタイルを身につけたい方に最適です。

【書誌 ID= 2004594208】配架場所: 理工学図-東館 2F 図書 007.64||MAT

##### ■ [Python によるデータ分析入門 : pandas、NumPy、Jupyter を使ったデータ処理 / Wes McKinney 著 ;瀬戸山雅人, 小林儀匡訳 \(第 3 版\)](#)

pandas 開発者による定番のデータ分析入門書。NumPy、Matplotlib、Jupyter Notebook など、科学技術計算に不可欠なライブラリを総合的に扱い、実際のコードを通してデータ操作と可視化を学べます。金融・統計・機械学習など幅広い分野に応用可能で、最新版 pandas 2.0 にも対応しています。

【書誌 ID= 2004570489】配架場所: 理工学図-東館 2F 図書 007.64||MCK

## ■ [scikit-learn、Keras、TensorFlow による実践機械学習 / Aurélien Géron 著；長尾高弘訳（第3版）](#)

Python を代表する機械学習ライブラリである scikit-learn、Keras、TensorFlow を組み合わせ、幅広い分野の機械学習手法を体系的に解説した定番書です。回帰・分類から深層学習、自然言語処理までを実践的に学べる内容に加え、第3版では大規模言語モデル(LLM)や拡散モデルなどの生成AIの基礎も網羅しています。ノートブック形式のサンプルコードを通して、理論と実装の両面から実践的スキルを習得できます。

【書誌 ID= 2004593838】配架場所：理工学図-東館 2F 図書 007.13||GER

## ■ [Python ではじめる機械学習 : scikit-learn で学ぶ特徴量エンジニアリングと機械学習の基礎 / アンドレアス・C・ミュラー、サラ・グイド著；中田秀基訳](#)

scikit-learn 開発に携わる著者が、機械学習の基礎から特徴量エンジニアリング、モデル評価・改善までを丁寧に解説した入門書です。ニューラルネットワークの前段階として、機械学習の原理と実装の流れを理解したい方に適しています。図表や実例を用いながら、データ前処理やモデル選択など実務で重要な要素を学ぶことができます。

【書誌 ID= 2004429082】配架場所：総合図-A棟 3階 学習用図書 007.13||MUL

## ■ [Python と Keras によるディープラーニング / François Chollet 著；クイープ訳](#)

ディープラーニングにより、画像認識や自動翻訳の精度は飛躍的に向上し、現代のソフトウェア開発者にとって必須のスキルとなりつつあります。Keras は TensorFlow 上で動作する Python 製の高水準 API であり、数学やデータサイエンスの専門知識がなくてもディープラーニングを直感的に実装できます。Keras の開発者が著した本書は、初心者から経験者までを対象に、豊富な例題と図解を通してディープラーニングの理論と実装を学べる定番書です。画像分類、時系列予測、テキスト生成などの主要トピックを実践的にカバーしています。

【書誌 ID= 2004452459】配架場所：理工学図-東館 2F 図書 007.13||CHO

### 3-2. 有用なナビゲートツール: Web ページ

- Python 公式ドキュメント(<https://docs.python.org/ja/3/>)：文法・標準ライブラリの一次情報。
- Google Colab(<https://colab.google/>)：環境構築不要で Notebook(ブラウザ上でコードと説明文と一緒に書ける形式)を実行。必要に応じて機械学習の学習を速くできる GPU も利用可。
- Kaggle(<https://www.kaggle.com/>)：データセット + Notebook で実践学習。
- Udemy(<https://www.udemy.com/>)：動画講座で基礎から応用まで体系的に学べる。

#### 本文中で紹介している図書・雑誌について

図書名・雑誌名の後ろに「書誌 ID」(10桁の数字)の記載があるものは大阪大学で所蔵しています。この10桁の数字で大阪大学 OPAC(蔵書検索システム)が検索できます。

