

やさしい

AIプログラミング 入門

Pythonで学ぶAIの仕組みと応用

日向俊二●著



■ サンプルファイルのダウンロードについて

本書掲載のサンプルファイルは、下記 URL からダウンロードできます。

<https://----->

- 本書の内容についてのご意見、ご質問は、お名前、ご連絡先を明記のうえ、小社出版部宛文書（郵送または E-mail）でお送りください。
- 電話によるお問い合わせはお受けできません。
- 本書の解説範囲を越える内容のご質問や、本書の内容と無関係なご質問にはお答えできません。
- 匿名のフリーメールアドレスからのお問い合わせには返信しかねます。

本書で取り上げられているシステム名／製品名は、一般に開発各社の登録商標／商品名です。本書では、™ および ® マークは明記していません。本書に掲載されている団体／商品に対して、その商標権を侵害する意図は一切ありません。本書で紹介している URL や各サイトの内容は変更される場合があります。

はじめに

現代社会では、AI はさまざまなところで使われています。スマホで撮った写真を加工したり、会話を外国語に翻訳したり、イメージや動画を自動的に生成したりすることが日常的に行われるようになりましたが、これらはすべて背後で AI が機能しています。サイトや動画を見るときに、あなたの嗜好や過去の閲覧傾向からおすすめの品物や動画などが表示されるのも、背後で AI が活躍しているからです。

本書では、最初に AI の概要について説明したあとで、AI で扱うデータについてと、AI が機能するしくみをきわめて簡素化単純化した例で示します。そのあとで自由に利用できる AI ライブラリを応用するいくつかのプログラムを説明します。

AI では一般に膨大なデータを複雑なアルゴリズムを多数使って目的を達成しようとします。そのためには、AI で処理できるようにデータを加工したうえで、難解な数式を理解してモデルを構築する必要があります。そのためのそれぞれの領域はとても深いので、それらをすべてゼロから自分一人で行うことは不可能です。しかし、多くの研究開発者が作り上げてきた既存のライブラリを活用することで、プログラマは AI を容易に活用することができます。

本書の前半では、複雑なデータ処理を簡略化し、プログラムとして実装するために必要な最小限のコードで AI の仕組みを説明します。いいかえると、あくまでも AI についてざっくりと、とはいえ、できるだけ具体的に理解することを目標として、可能な限り単純化して重要な点に焦点を当てます。

本書の後半では、便利で使いやすい既存の AI ライブラリを活用する方法を示します。

本書ではこれらの説明にプログラミング言語として Python と、Python から利用できるさまざまなライブラリを使います。そのため、本書に記載した範囲で Python やライブラリの使い方に関連することがらをひと通り説明します。ですから、Python になじみがない場合でも本書の内容を理解してプログラムを実行してみることができますが、Python についてより多くの知識があるとさらに理解しやすいでしょう。

AI の用語、概念、そしてそこで使われる数式やプログラムコードの深い理解は難しい
うえにその範囲は広範なので、AI の専門書をひもとくのは大きなチャレンジになるでし
ょう。しかし、本書で基本的な言葉や概念、プログラムでの扱い方を理解することで、
広大な AI の世界への道が開かれるでしょう。

2024 年初夏 著者しるす

本書の表記

- > Windows のコマンドプロンプトを表します。Linux や WSL など UNIX 系 OS の場合は一般的にコマンドプロンプトとして \$ が使われます。
- >>> Python のインタラクティブシェルのプロンプトを表します。
- abc** OS や Python のプロンプトに対してキーボードから入力する部分を太字で示します。
- abc* 斜体文字は、そこに具体的な文字や数値が入ることを表します。たとえば「python -m pip install *module*」は、*module* には numpy などの具体的なモジュール名が入ります。
- ... 書式の説明において任意の個数を記述できることを示します。
- [] 書式の説明において省略可能な部分であることを示します。
- () たとえば、「random という名前の関数」や「関数 random」と表記する代わりに、単に「random()」と表記することがあります。



本文を補足するような説明や、知っておくとよい話題です。

ご注意

- 本書の内容は本書執筆時の状態で記述しています。Python や AI ライブラリは、バージョンによって異なる点があり、本書の記述と実際とが異なる結果となる可能性があります。
- 本書は Python や AI のすべてのことについて完全に解説するものではありません。必要に応じて Python やその他のドキュメントなどを参照してください。
- 本書のサンプルは、プログラミングを理解するために掲載するものです。実用的なアプリとして提供するものではありませんので、ユーザーのエラーへの対処やセキュリティ、その他の面で省略してあるところがあります。

本書に関するお問い合わせについて

本書に関するお問い合わせは、sales@cutt.co.jp にメールでご連絡ください。

なお、お問い合わせは本書に記述されている範囲に限らせていただきます。特定の環境や特定の目的に対するお問い合わせ等にはお答えできませんので、あらかじめご了承ください。特に、特定の環境におけるインストールや環境設定、使い方、読者固有の環境におけるエラーなどについてお問い合わせいただいてもお答えできませんのでご了承ください。

お問い合わせの際には下記事項を明記していただきますようお願いいたします。

- 氏名
- 連絡先メールアドレス
- 書名
- 記載ページ
- お問い合わせ内容
- 実行環境

目次

はじめに	iii
------------	-----

第 1 章 AI の概要 1

1.1 AI とは.....	2
◆ AI の定義	2
◆ AI と AI でないもの	2
◆ AI の特性	4
◆ AI ができること	7
1.2 IPO	8
◆ 従来の IPO	8
◆ AI の IPO	8
1.3 機械学習	10
◆ 教師あり学習	10
◆ 教師なし学習	10
◆ 強化学習	11
◆ データとモデル	11
練習問題	12

第 2 章 Python の基礎 13

2.1 Python との対話	14
◆ インタラクティブシェル	14
◆ プロンプト	15
◆ 単純な加算	15
◆ print() を使った出力	17
◆ 行継続文字	18
2.2 スクリプトファイル.....	20
◆ ファイルの作成	20
◆ ファイルの保存	21
◆ スクリプトの実行	21
◆ 実行方法の比較	22

2.3 Python の主なシンタックス.....	24
◆ コメント	24
◆ for	24
◆ range	26
◆ while	27
◆ if	30
◆ else	33
◆ elif	33
◆ and、or、not	34
◆ import	35
◆ __version__	35
◆ 演算子	36
2.4 モジュールの利用.....	38
◆ モジュールのインストール	38
◆ import	39
◆ モジュールの属性へのアクセス	39
◆ random モジュール	40
練習問題	42

第 3 章 AI のデータ43

3.1 情報とデータ	44
◆ 有意義なデータ	44
◆ 極端なデータ	45
3.2 プログラムの中のデータ	46
◆ Python でのデータの表現	46
◆ 2 次元データ	48
3.3 データの可視化	50
◆ 折れ線グラフ	50
◆ 散布図	53
◆ ヒストグラム	55
3.4 データの基本的な値.....	57
◆ 基本統計量	57
◆ 平均値	58
◆ 加重平均	59
◆ 最大値と最小値	60
◆ 中央値	60
◆ 最頻値	61

◆ 分散	62
◆ 標準偏差	63
◆ 極端なデータの除外	64
◆ 正規化	65
3.5 データの生成	67
◆ CSV ファイル	67
◆ データの作成例	68
練習問題	73

第 4 章 AI のしくみ75

4.1 予測	76
◆ 単回帰と重回帰	76
◆ 仕組みの理解とデータ	76
4.2 単回帰	77
◆ 直線回帰	77
◆ 多項式回帰	82
4.3 重回帰	87
◆ 重回帰とは	87
◆ 重回帰のプログラム例	87
4.4 AI の仕組み	92
◆ AI の考え方	92
練習問題	94

第 5 章 Streamlit95

5.1 streamlit の基礎	96
◆ 概要	96
◆ streamlit のインストール	96
◆ シンプルなサンプル	96
5.2 streamlit のウィジェット	101
◆ コマンドボタン	101
◆ 数値入力	102
◆ テキスト入力	103
◆ テキストエリア	105
◆ ファイルアップローダー	106

◆ マルチセレクト	106
◆ 散布図	108

練習問題	110
-------------	-----

第 6 章 予測111

6.1 単回帰分析アプリ	112
◆ 2 次曲線による回帰	112
6.2 重回帰分析アプリ	115
◆ アプリのイメージ	115
◆ モジュールのインポート	117
◆ ページの UI	118
◆ CSV ファイルの読み込み	120
◆ グラフの表示	121
◆ 分析と予測	121

練習問題	126
-------------	-----

第 7 章 画像認識127

7.1 画像認識の流れ	128
◆ 画像認識	128
◆ 認識モジュール	130
◆ 画像モジュール	133
7.2 画像認識 Web アプリ	134
◆ アプリの概要	134
◆ プログラム	136

練習問題	142
-------------	-----

第 8 章 自然言語認識143

8.1 テキストの操作	144
◆ 自然言語の認識	144
◆ 使用するモデル	146
◆ 英語のモデル	149

8.2 自然言語認識アプリ	151
◆ アプリの概要	151
◆ プログラム	152
練習問題	157

付録159

付録 A コマンド	160
付録 B トラブル対策	163
付録 C 練習問題解答例	168
付録 D 参考リソース	176
索引	177

第 1 章

AI の概要

ここでは AI についてその考え方や大まかな使われ方などについて概説します。この章ではさまざまな言葉が登場しますが、AI が具体的にどのようなものであるかという点に焦点を当ててください。

1.1 AIとは

AI (Artificial Intelligence) は、日本語では一般的には人工知能と訳されます。

◆ AIの定義 ◆

さまざまな人や組織などがAIの定義を試みっていますが、現在のところ、AIの厳密な定義は定まっていません。本書全体を通じて説明しますが、AIは曖昧なものを曖昧なものとして扱います。また、扱うのが限られた範囲ではなく、無限ともいえます。そのため、AIの厳密な定義は従来の方法では不可能なのかもしれません。しかし、AIを扱う以上、AIがどのようなものであるのかということは最初に知っておく必要があるでしょう。

◆ AIとAIでないもの ◆

AIというものを、AIでなければ扱えないものと、AIとは関係なくAI登場以前の方法だけで扱えるものにわけて考えることによって、AIというものをより明確にしましょう。

最初の例は果物のみかんの選別です。

収穫した多数のみかんを重さでS、M、L、LL、LLLに分けるとします。このときには、重さを計って分けて行けば良いので、AIとは関係なく分けることができます。つまり、単に重さを測定してSの範囲の重さのものはSに、Mの範囲の重さのものはMに……、と分ければよいこととなります。したがって、重さという数値だけで分類することができるといえます。

次に、みかんの品質で次のように分けるとします。

- A級品：形が整っていて、表面に傷や変色がほぼない良いみかん。
- B級品：形や表面に些細な傷や変色が見られるみかん。
- C級品：形がいびつか、傷や変色が多いみかん。

この場合、カメラでみかんの画像を取得することは容易ですが、その画像から「形が整っている」かどうかを簡単に判断する方法はありません。「整った形のみかん」は完全

な球体のような形でも球体を上下から少し押しつぶした形でもありませんから、そもそも「整った形」をコンピューターで理解できる方法（すなわち数値）で表現することさえ困難です。

みかんの表面の傷や変色についても、些細な変色なのか、味には影響のない傷なのか、それとも中で腐っていてそれが表面にまで及んでいるための変色なのか、みかんの画像から判断するのはとても困難です。このように、AIを使わない従来のコンピューターの方法では判断することは事実上不可能といえるでしょう。

人間は多数のみかんを見る経験を積むことで、良いみかん、普通のみかん、ダメなみかんと判断することができるようになります。それと同じように、コンピューターでみかんの品質を判断するには、「良いみかん」を知るための学習を行って、ミカンをコンピューターが理解できる方法で選別するためのアルゴリズム（モデル）を作ることが必要になります。このような場合に AI が必要とされます。

もうひとつの例を考えてみましょう。

犬を写したカラー写真があるとします。コンピューターで写真を扱うときには、多数の小さなピクセル（画素）で構成されているイメージ（画像）として扱います。カラーイメージの場合、各ピクセルは赤、緑、青の色の強さの値（RGB 値）で表現されます。たとえば、各色の強さを 0～255 で表すとすると、RGB 値が (255, 255, 255) ならばそのピクセルの色は白、RGB 値が (255, 0, 0) ならば赤、RGB 値が (0, 0, 0) ならば黒になります。

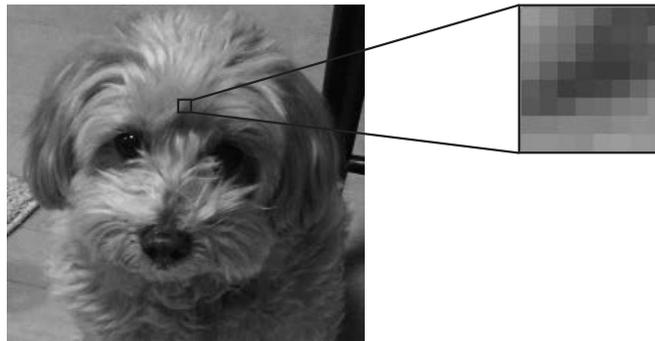


図1.1 ●イメージと部分を拡大した状態

このカラー写真を白黒に変換するには、各ピクセルの RGB 値を加算して 3 で割ることで白黒の階調（グレイ）の値にします。あるピクセルの RGB のそれぞれの値を r 、 g 、 b

とすると、白黒の濃さの値 v は次の式で求めることができます。

$$v = (r + g + b) / 3.0$$

この操作は、イメージを構成している各ピクセルの RGB 値から単純な計算で行うことができます。つまり、カラー画像をモノクロ画像に変換するような作業には AI は必要ありません。

一方、犬を写したイメージから犬の部分だけを取り出して、背景を消したいとします。コンピューターはそれぞれのピクセルの色の情報だけでは、そのピクセルが犬を表現しているピクセルなのか背景を表現しているピクセルのかわかりません。犬を背景と区別するには、たくさんの犬のイメージから犬を表す特徴を調べて、背景と区別することができるようになるための学習が必要になります。さらに、画像のその部分が犬であることをコンピューターが理解できる方法で判定するためのアルゴリズム（モデル）を作ることも必要になります。このような場合に、AI が必要とされます。

◆ AI の特性

次に、現在 AI が利用されているケースを少し詳しく見てみましょう。そうすることによって、AI がどのようなものであるかということをもより深く理解できるでしょう。

ウェブサイトや動画を見ていると、おすすめの品物やおすすめの動画などが表示されることがあります。AI はおすすめの品物や動画などをランダムに表示しているのではなく、ユーザーの過去の閲覧傾向から AI がそのユーザーの嗜好や傾向を予測して提示します。これは、ユーザーの過去の操作から AI が学習していることを意味します。ときには的を外れた「おすすめ」が提示されることもありますが、その理由は、過去の閲覧履歴が十分でないこと（学習データが十分でないこと）や、そもそも AI は正解を示すためのものではなく、正しい可能性が高いと思われるものを示すものだからです。

イメージや動画を自動的に生成したりする AI は、たとえば「ヨットが浮かんでいる海岸風景」などと要求すると、要求に従ったイメージや動画を作ってくれます。しかし、これも何も無いところから生み出されるのではなく、ヨットや海岸の風景の画像や動画をたくさん学習したあとで、複雑なアルゴリズムを使って、提示された問題にふさわしいと思われる画像や動画を生成します。

そして、ほとんどの場合、ただ一つの画像や動画ではなくて、いくつかの候補を提示します（提示された問題の解答として適切である確率が高いものから順に提示することがよくあります）。図 1.2 に示すのは、Bing Image Creator という Microsoft が提供している画像生成 AI ツールで、「浜辺で夕陽を見ている少年」というお題で画像を生成した例です。



図1.2●Bing Image Creatorの生成画像の例

生成された4種類の画像は、「浜辺で夕陽を見ている少年」という文に関連が深いであ

ろうという観点から計算された画像として Bing Image Creator が提示したものです。この画像のどれが正解というものではありませんし、画像を要求したユーザーが想像していたものとはどれもかなり違うかもしれません。

アプリによっては可能性が一番高いものを一つだけ提示することがありますが、いずれにしても、AI が生成する答えはひとつではなく、提示された問題に対する答えとして適切である可能性が高いものを提示してくることを意味します。

AI で画像から人間や動物などを検出することも行われています。その場合、AI のシステムは人間や特定の動物の画像をたくさん読み込んで、それぞれの特徴を持つものを人間であるとか犬であるとか判断します。このとき、画像認識で「犬」と判定されたとしても、正確には「犬である可能性が 98.5%、キツネである可能性が 1.5%」というような確率として得られるだけで、絶対に「犬」であると断定されるわけではありません。

AI が会話を自然な外国語に翻訳するときにも、単に単語や熟語、典型的な言い回しを他の言語のそれに置き換えているのではなく、たくさんの自然な文を学習して、目的の言語で自然であると思われる可能性が高い表現を生成して提示します。

これらのことからわかる重要な点が二つあります。一つは AI が答えを出すためには、得られている情報から答えを導く方法を学習する必要があるということです。もう一つは、得られる答えは一つではなく、適切である可能性が高いものを AI が提示するという点です。つまり、AI の結果は正しい答えではなく、正しい答えである確率が高いものであるということが出来ます。

AI が登場する以前のコンピューターの目的は、原則として、ただ一つの正解または限りなく正解に近いもの（円周率の近似値など）を求めることでした。しかし、AI では必ずしも正解のない問題に対する解答を求めることがあるという点で従来のコンピューターの考え方とはまったく異なります。



ChatGPT でのチャットには「ChatGPT は間違いを犯すことがあります。重要な情報は確認をお考えください。」というメッセージが表示されます。これは ChatGPT が返す回答が唯一の正解でないことを示しています。

◆ AI ができること

現時点で、実用化されている AI システムのほとんどは、自然言語処理や音声・画像の認識、文章や画像の生成、特定の分野の値の予測など、特定の分野を対象とするものです。天気を予測するシステムで絵を描くことはできませんし、文章を他言語に翻訳するシステムで作曲することはできません。ChatGPT のように、さまざまな問題に対する解答を返したり、翻訳やプログラムコードの生成などを行うといった、かなり多方面で使うことができるものもありますが、どんな問題でも解決するようなものは今のところ存在していません。

とはいえ、AI を構成する技術はどんどん進歩しているので、今後状況はどんどん変わっていくことが予想されます。

1.2 IPO

IPOとは、Input（入力）、Process（作業や処理）、Output（出力）を表します。

◆ 従来のIPO ◆

従来の一般的なコンピュータプログラムは、なんらかの入力を処理してその結果を出力します。



図1.3●Input-Process-Output

たとえば、体重と身長から求まる肥満度を表す体格指数であるBMI（Body Mass Index）を求めるプログラムでは、入力は体重（kg）と身長（m）で、出力はBMI値であり、次の式でBMIを計算することがプログラムの処理になります。

$$\text{BMI} = \text{体重} \div (\text{身長} \times \text{身長})$$

このとき、入力である特定の体重と身長の値から導かれる答えは必ず一意（ただ一つに定まる値）であり、他の値である可能性はゼロです。言い換えると、従来のIPOの出力は常に正解だけです。

◆ AIのIPO ◆

AIで行うことも基本的にはInput-Process-Outputですが、従来のコンピュータのプログラムで行うのとは異なる点があります。

まず、Inputについては、通常、AIでは多数のデータを入力とします。たとえば、年齢から身長を予測する場合、モデル（アルゴリズム）を決めるために多くの人の年齢と身長のデータ（学習用のデータ）を入力します。そして、特定の年齢のヒトの身長を予

測するために、予測したい人の年齢を入力して、その年齢のヒトの身長として可能性が高い身長を結果として出力します。

AI の Process では、学習用のデータからモデル（アルゴリズム）を決めて、求めたいデータからいくつかの結果の候補を計算します。得られるデータによって実際に行うことが違うという点で、あらかじめ決まったパラメーターで決まったアルゴリズムを実行する従来の AI ではない処理とは大きく違います。

すでに説明したように、AI の Output では、可能性が高い候補が複数出力されます。システムによってはその中の最も確率の高いものを結果として一つだけ出力する場合がありますが、その結果が 100% 正しいという保証はありません。これが AI の出力の大きな特徴です。



.....

入力データは、目的と使い方によって、テストデータ、学習データ、訓練データ、目的データ、教師データ、検証データなどと呼ぶことがあります。それぞれの定義は異なりますが、この段階では処理に必要なさまざまなデータ全体が入力データであると理解しておいてください。入力データの用語に限らず、AI では目的や使い方、そして使われるアルゴリズムやコミュニティによって用語が異なることがよくあります。また、AI は現在も進化しているので、新しい用語が使われたり、用語の意味が変わったりすることがあります。

.....

1.3 機械学習

AIでは、多数のデータから学習して出力を生成します。これを機械学習（machine learning）と呼ぶことがあります。機械学習は、一般に、教師あり学習、教師なし学習、強化学習に分類されます。

◆ 教師あり学習

教師あり学習（Supervised Learning）は、入力に対して適切な出力を求める問題です。学習するために、問題と正解を与え、正解とデータの特徴量の関係性を調べて、予測値を正解に近づけることを目標にします。

教師あり学習では、一般に、回帰問題と分類問題とを扱います。

回帰問題は、たとえば過去の売上げから将来の売上を予測したり、多数のヒトの年齢と身長とのデータを与えて、特定の年齢に対して期待される身長を予測します。

分類問題は、それぞれのデータが所属するカテゴリを推定します。たとえば、動物の画像を入力データとして、犬であるか猫であるかを識別します。

◆ 教師なし学習

教師なし学習（Unsupervised Learning）は、入力の特徴を求める問題です。正解はなく、学習のために問題を入力し、共通する特徴を見つけます。

教師なし学習では、一般に、クラスタリングと次元削減とを扱います。

クラスタリングは、データをいくつかのクラスタに分けます。たとえば、検索履歴から興味ごとのグループに分けます。

次元削減は、データを低い次元に圧縮することで重要な情報を際立たせます。たとえば、さまざまなパラメーターが関係するデータの主成分を分析します。

◆ 強化学習

強化学習 (Reinforcement Learning) は、明らかな正解がない場合に行われ、最良と思われる結果を求める問題です。

たとえば、将棋の手には非常に多くのバリエーションがあって正解はありません。そのような場合に、最大の報酬 (勝利) を得られるように、最適な行動を学習します。



Note

教師あり学習とか教師なし学習という日本語が一般に使われていますが、これは英語を翻訳した訳語です。少々不自然と感ずるかもしれませんが、このような呼び方が慣例なので覚えておきましょう。

◆ データとモデル

機械学習では、適切な多数のデータの特徴や結果との相関性をみつけて結果を導きます。そのため、問題の解答を機械学習で得るためには、最初に適切なデータを多数集めることが必要になります。単にやみくもに集めたデータそのままでは、たいていの場合、役に立ちません。そのため、データの処理や加工が必要になります。データについて詳しくは第3章「AIのデータ」で取り上げます。

結果を導くためのモデル (アルゴリズム) は、問題の種類とそれに対するアプローチの種類によっていろいろあります。モデルについては、第4章「AIのしくみ」で具体的例を示してより詳しく取り上げます。

練習問題

問題 1-1

次の中から誤っているものを選んでください。

1. AI では常に正しい解答が得られる。
2. AI が示す出力は適切である確率が高い。
3. AI では結果を得るためにたくさんのデータを使うことが普通である。
4. どんな問題でも解決できる AI システムは現在のところない。

問題 1-2

次の中から誤っているものを選んでください。

1. 教師あり学習は問題と正解を与えて、入力に対して適切な出力を求める問題である。
2. 教師なし学習は、正解はなく入力を与えて入力の特徴を求める問題である。
3. 強化学習は、明らかな正解がない事項の最良と思われる結果を求める問題である。
4. AI の学習は、少ないデータからでも常に最良の結果を求めることができる。