

# 体感しながら学ぶ AI とデータサイエンス入門授業の試み

## —Wolfram 言語によるプログラミング実践—

金光安芸子\*1・北村美穂子\*2・吉田智子\*2

Email: akikok@wolfram.com

\*1: Wolfram Research, Inc.

\*2: 京都ノートルダム女子大学 社会情報課程

◎Key Words AI・データサイエンス教育, プログラミング実習, Wolfram 言語, 教材開発

### 1. はじめに

どのような分野においても AI やデータサイエンスの知識が不可欠な時代である。その流れを受けて、2021 年度から文部科学省は「数理・データサイエンス・AI 教育」を体系的に行なっている大学等に対して認定制度<sup>1)</sup>を開始し、文系理系を問わず全ての学生にこの教育を提供するよう推進している。

筆者らが授業を行なった京都ノートルダム女子大学は、文系女子大学でありながら全学的な情報教育を数年前から積極的に行っており、2021 年度に上記の認定制度（リテラシーレベル）の認定を受けた。特に文系学部 of 学生や女子学生にとって、データサイエンス教育を含む情報教育における教材の選び方は非常に重要である<sup>2)</sup>。身近な題材で学ぶ必要性を実感でき、学生自身のデザイン力や創造性を発揮できる教材が受け入れられ易い。

また、データサイエンスの授業において、講義で学ぶ内容をプログラミング実習で体感することは、仕組みを理解するために非常に重要である。ただ、プログラミング言語に馴染みのない学生にとって、言語を学ぶこと自体が負担になることは望ましくない。

本報告では、これらの条件を満たすものとして京都ノートルダム女子大学の「データ活用プログラム」の必修科目の一つである「AI とデータサイエンス入門」で実践した「Wolfram 言語<sup>1)</sup>によるプログラミング実習」について紹介する。

### 2. 「AI とデータサイエンス入門」の授業

#### 2.1 授業の概要

「AI とデータサイエンス入門」（2 単位、90 分授業 15 回）は、2021 年度以降に入学した 2 年生以上の全学部生が選択できるもので、北村と金光の 2 名が 2 クラス（計 21 名）を担当する形で 2022 年度の後期授業（2022 年 9 月～2023 年 3 月）で実践した。すべての授業は原則「講義＋Wolfram 言語によるプログラミング実習」という形式で行った。この授業全体の方針、内容および実践結果については、2023PC カンファレンス分科会の「データサイエンス」において「初学者のための AI・データサイエンス入門授業の試み」で詳しく発表する。

#### 2.2 Wolfram 言語の実習環境

Wolfram 言語は、計算知識エンジン Wolfram|Alpha<sup>2)</sup>や数式処理システムとして 30 年以上の歴史のある Mathematica<sup>3)</sup> のベースとなっているプログラミング言語である。微積分、線形代数、統計などの数学関数はもちろん、画像分析や音声分析、自然言語処理や機械学習など AI とデータサイエンスに欠かせない高度な関数が豊富に揃っているため、講義の内容に沿った実習が可能である。

本実習では、この Wolfram 言語の実行環境として、Wolfram Cloud<sup>4)</sup> を利用した。これは、Web ブラウザからアクセスでき、Wolfram ID というアカウントを作成してサインインすると利用できる（図 1）。PC にアプリケーションをインストールする必要がないため、大学の PC 教室でなくとも、学生が所有する PC、モバイルデバイス、スマートフォンなどでどこからでも同じ環境でプログラミングができる。

実習用の教材テキストは Wolfram Notebook で作成し、学生には Wolfram Cloud 上で公開した。Wolfram Cloud では、ファイルのシェアや公開が簡単にできるため、教材の配布や学生の課題提出、質問のやりとりなど、教員と学生のコミュニケーションも双方向でスムーズに行うことができる。

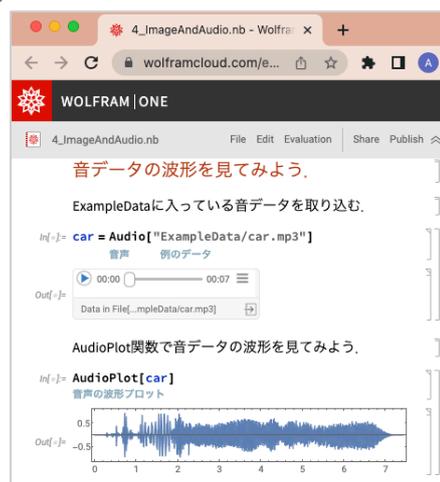


図 1 Web ブラウザで利用できる Wolfram Cloud

#### 3. Wolfram 言語を使った実習の実践例と学生の反応

全 15 回の授業を「講義＋Wolfram 言語によるプログラミング実習」という形式で行ったが、ここでは、学生が特に意欲的に取り組んだ 3 つの実習内容について紹介する。

<sup>1</sup> <https://www.wolfram.com/language/>

<sup>2</sup> <https://www.wolframalpha.com/>

<sup>3</sup> <https://www.wolfram.com/mathematica/>

<sup>4</sup> <https://www.wolframcloud.com/>

### 3.1 アートプログラミング

「プログラミング言語概論」の回(第6回)では、アルゴリズムやデータ構造、フローチャート、プログラミング言語の種類や特徴を講義で学んだ後、条件分岐や反復処理、変数や関数の作成のプログラミング実習を行った。第7回では、それらを組み合わせて大小様々な円のグラフィックスを描くアートプログラミングの課題に取り組んだ。

実習での基本となるプログラムを図2に示す。Circle関数を使って円を一つ描く短いコードから始まり、Table関数を使って複数の円を描く方法、RandomReal関数やRandomColor関数を組み合わせてランダムな大きさや色でアートを作る方法を学んだ。学生は、教材ファイルに記載してあるプログラムをまずはそのまま実行し、プログラムの意味を考えながら数字や関数を少しずつ変えてオリジナルのアート作品に仕上げた。

さらに応用編として、Circle関数をDisk関数で円盤に、Polygon関数で多角形に、Opacity関数を追加して透明度を与えるなどの方法を提示したところ、大半の学生が意欲的に挑戦し、工夫してアートプログラミングの作品作りに取り組んだ。学生の作品を図3に示す。

学生からは、「自分の思うような色を付けるために試行錯誤するのが楽しかった」、「プログラミングは難しいと思っていたが、実際自分でやってみて様々な色や形が表現出来るのがとても面白かった」、「自分の創造性で多くの芸術作品を作ることができてよかった」など、学生のやる気を示す感想が多く見られた。

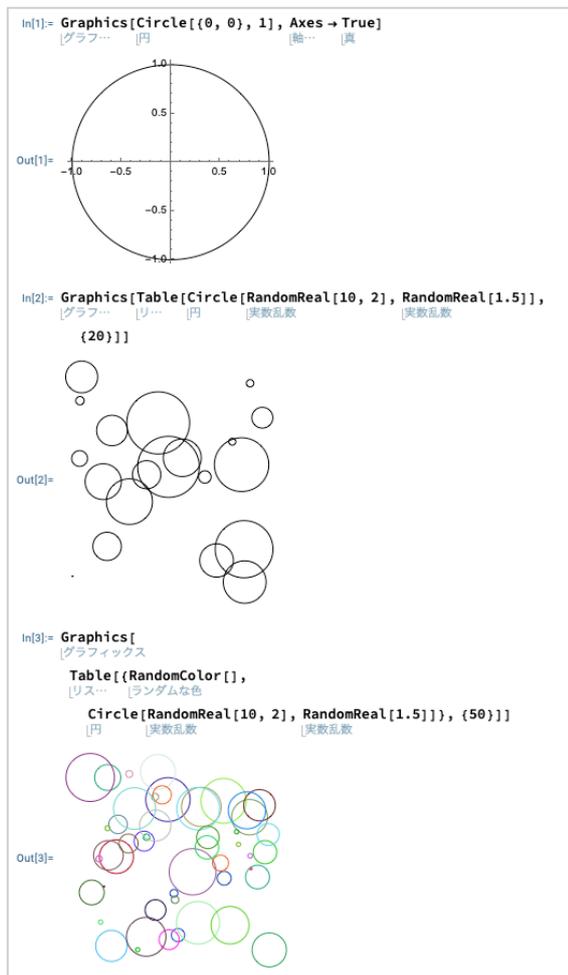


図2 アートプログラミングの基本プログラム

### 3.2 文学作品のワードクラウド

「データビジュアライゼーション」の回(第12回)では、可視化の種類(棒グラフや地図上の可視化など)、可視化の重要性、チャートジャンクについて講義で学んだ後、ワードクラウドで可視化するプログラミング実習を行い、可視化の有効性について共に考えた。

ワードクラウドは、アンケートや文章などのテキストデータから出現回数の多い単語をより大きく表示することで、どのようなワードがより多く出てきているのかを瞬時に認識することができるデータビジュアライゼーションの一つである。ニュースや新聞、TVの情報番組などでもよく利用されており、学生にも馴染みがある。

実習では、まず授業アンケートのテキストファイルをWolframCloudにアップロードした後、アップロードしたファイルをImport関数で読み込み、WordCloud関数でワードクラウドを作成した。さらに、ハート型や幾何図形を指定して異なる形のワードクラウドを作成した。図4にWolfram言語での具体的な手順を示す。

ここまでの実習を全員で行った後、教員がレポート課題について説明した。課題の内容は、表1の10個の文学作品から2作品を選んでワードクラウドを作り、そこからわかることや工夫した点を記述することである。

表1 ワードクラウドの課題とした文学作品

「吾輩は猫である」(夏目漱石)
「人間失格」(太宰治)
「みだれ髪」(与謝野晶子)
「源氏物語 桐壺」(紫式部, 与謝野晶子訳)
「源氏物語 夕顔」(紫式部, 与謝野晶子訳)
「母を尋ねて三千里」(アミーチス)
「みにくいアヒルの子」(アンデルセン)
「白雪姫」(グリム)
「ラプンツェル」(グリム)
「若草物語」(オルコット)

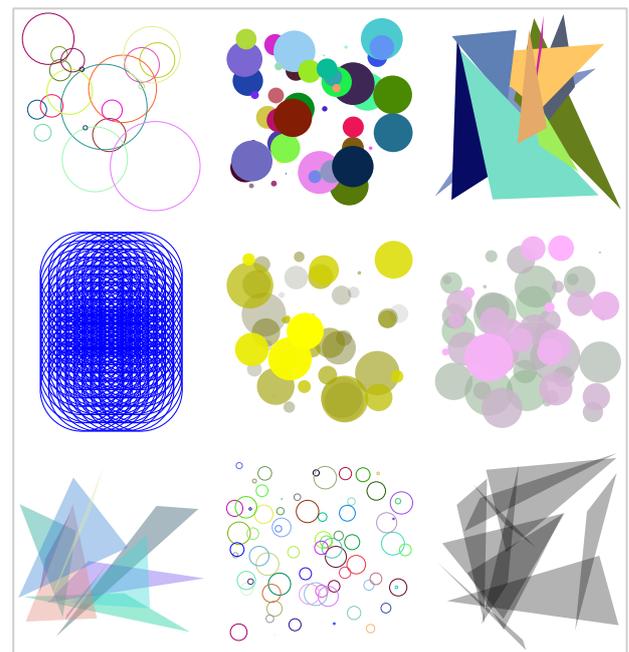


図3 学生が作成したアートプログラミングの作品

なお、レポートに使用する文学作品のデータは、教員が青空文庫<sup>5</sup>から入手した全文テキストをPythonのMeCab<sup>6</sup>を利用して形態素解析をした後、ストップワード（「を」「は」「です」などの助詞や助動詞）を除去してテキストデータファイルを用意した。

学生らがこのレポート課題に取り組んだのは冬休み期間中であったが、休み中にも関わらず、学生から「作品に合わせた形にしたいがどうしたらいいか?」、「リストにない文学作品でワードクラウドを作りたい」という質問や要望があり、意欲的に取り組もうとする様子が伺えた。図5は、学生の要望を受けて追加でWeb上に公開したオリジナルの形のワードクラウドを作る手順である。



図4 ワードクラウドの作成手順

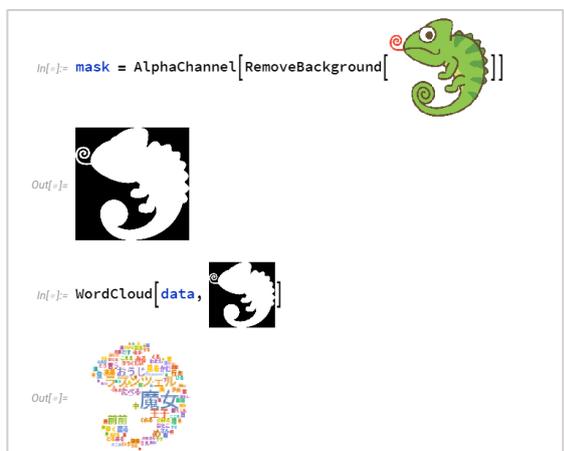


図5 好きな画像からワードクラウドを作成する方法

学生らはそれぞれの文学作品のイメージや登場人物に合わせて色や形を工夫し、バリエーションに富んだワードクラウドを作成した。学生らは、課題として指示された2作品以上のワードクラウドを作成したり、自らWolfram言語のオンラインドキュメントを調べて色や配置を変化させたりしており、筆者らの期待以上に意欲的に取り組んでいた。図6にレポート課題で提出された学生らのワードクラウドの作品の一部を紹介する。

また、課題の「ワードクラウドからわかること」として、表2のような学生の意見があった。文学作品をワードクラウドで可視化することによって新たな発見があったことがわかる。

表2 ワードクラウドからわかること

「主人公の名前や登場するキャラクターのワードが大きく表示され、それらのワードが多く登場していることがわかった。」
「すべての小説において、一人称、二人称、人物名がよく出てくるということがわかった。また、動詞も多く、特に多いのは“いう”という動詞であることがわかった。このことから、人物が何を言ったのかという文章が多いのではないかとと思われる。」
『ラプンツェル』のワードクラウドからは、主人公の“ラプンツェル”よりも“魔女”の方が多く出現していることに驚いた。この結果は、ワードクラウドを作成してみないと分からないことなので、これこそがワードクラウドの良さだと思った。」



図6 学生が作成した文学作品のワードクラウド

<sup>5</sup> <https://www.aozora.gr.jp/>  
<sup>6</sup> <https://taku910.github.io/mecab/>

### 3.3 機械学習でお菓子を見分けるアプリの作成

「機械学習」の回(第13回)では、機械学習の種類(教師あり学習, 教師なし学習, 強化学習)と教師あり学習の簡単な仕組みについて講義で学んだ後, 教師あり学習で2種類のチョコのお菓子を画像から見分けるプログラムを自作する実習を行った。

実習では, まず教員が用意した2種類のお菓子の画像データのうち, 訓練用の画像データにラベリングし, その訓練データを使って Classify 関数で学習させて判別器を作った。次にその判別器を使ってテストデータで判定実験を行い, 性能を調べた。さらに, 応用として Method オプションを使って異なる機械学習手法の判別器を作り, 判定結果を比較した。図7に教材テキストで示したプログラムの手順を示す。

学生らはこのプログラムを自ら実行して実験を行なった。そして, 判定結果やその根拠となる確率などから, 判定が正解あるいは不正解になった理由を考え, よりよい性能の判別器を作るにはどうしたらよいかを共に考えた。学生からは, 「訓練データに様々な角度での写真を増やすと良い」, 「訓練データの量を多くする」, 「データに適した手法を用いる必要がある」などの意見が聞かれた。

小さいプログラムでも身近な題材を用いて自分で実行し判定することによって, AIや機械学習を使ったサービスやアプリケーションが実際にどのような仕組みで動いているのかを学生自らが考え, 気づきを得たことが伺えた。

#### 1. 訓練データの準備

まず最初に訓練データを準備する。「きのこの山」と「たけのこの里」の画像にそれぞれラベルをつけて訓練データ `trainingdata` とする。

```
trainingdata = {  
    きのこの山 → "きのこの山",  
    たけのこの里 → "たけのこの里",  
    ...  
};
```

#### 2. 訓練データで機械学習の分類関数を学習させる

機械学習の分類関数 `Classify` に訓練データ `trainingdata` を使って学習させて「きのこの山とたけのこの里自動判別器」 `mytest1` を作る。

```
mytest1 = Classify[trainingdata]
```



#### 3. テストデータ (訓練データに使われていない画像) で判別器 `mytest1` の性能を確認する

テスト用の画像を直接 `mytest1` 関数の引数に入れて判別してみよう。

```
mytest1[きのこの山]
```

2番目の引数として, "Probabilities" を追加すると, 判定の根拠となる確率を見ることができる。上と同じ画像の判別の確率を見よう。

```
mytest1[きのこの山, "Probabilities"]
```

```
{ | きのこの山 → 2.15236 × 10-92, たけのこの里 → 1. | }
```

図7 お菓子を見分ける教師あり学習の手順

### 4. おわりに

本実習では, 講義で学習した知識の理解がプログラミング実習によって深まるよう, 題材選びを工夫した。特に, アートや文学作品, お菓子の判別アプリ作成など, 文系の学生にも受け入れられやすい身近な題材では, 「プログラミング=間違えずに打ち込む作業(写経)」に終わらず, 学生が自分の創造力を発揮し, 楽しく意欲的に取り組むことができた。

また, 比較的短いコードで様々なデータサイエンスのエッセンスを体感できる Wolfram 言語を毎回使用することにより, プログラム言語に不慣れな学生も徐々に実習や課題にスムーズに取り組むことができた。さらに Web 版の Wolfram Cloud 環境で大学が休みの期間でも自宅で作業が続けられたことも, 学生がやりたいと思ったタイミングを逃さず積極的に取り組めた要因の一つだと考える。

今回の実習テキストの一部は Wolfram Cloud 上に公開しており, だれでも利用することができる<sup>9)</sup>。すでに2023年度4月より華頂女子高等学校2年生を対象とした授業「メディア情報演習」にてこの教材の一部を活用した授業が実践されており, 高校生らがワードクラウドに取り組んでいる。

本教材は, 理系の専門ではない大学や小中高においても, AIとデータサイエンスをプログラミング的思考と共に楽しく学べる教材になりうると考える。

### 参考文献

- 文部科学省: “数理・データサイエンス・AI教育プログラム認定制度(リテラシーレベル)”, [https://www.mext.go.jp/a\\_menu/koutou/suuri\\_datascience\\_ai/00002.htm](https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm) (2019)
- 吉田智子, 有賀妙子, 真下武久: “ジェンダーインクルーシブなプログラミング教材の開発”, 情報処理, Vol62, No.12, pp.667-671 (2021).
- 金光安芸子, 北村美穂子, 吉田智子: “AIとデータサイエンス入門 - Wolfram 言語を使ったプログラミング実習テキスト”, <https://wolfr.am/ai-datascience-textbook> (2023)