

文理融合学部における AI 教育の実践： 文系にもわかりやすい実習を目指して

社会安全学部准教授 河野 和宏

1. はじめに

2019年に発表された AI 戦略2019^[1]を契機に、小学生から大学生までの各年代において、DS (Data Science : データサイエンス) や AI (Artificial Intelligence : 人工知能)、数理、プログラミングに関する教育が活発に行われるようになった。本稿が対象とする大学での教育に関しても、2021年に文科省により創設された「数理・データサイエンス・AI 教育プログラム認定制度」^[2]において、2021年8月時点で早くも78件¹⁾の教育プログラムが認定されるなど、大学における DS や AI に関する教育は現代において必須のものとなっている。関西大学でも、2021年度から「数理・データサイエンス・AI」に関する知識を習得できる全学共通カリキュラムを設置し^[3]、2021年度は2科目、2022年度はあらたに2科目追加し合計4科目の DS や AI に関する講義科目を提供している²⁾。

ここで、2021年度から本学で提供している2科目について詳しく見ると、この2つの科目は DS や AI に関する基礎知識を身に付け、社会でどのように活用しているかを知る教養科目となっている。どちらの科目も、受講生は、現代社会では DS・AI に関する知識・スキルは必須であると感じて意欲的に学習に取り組み、授業内容も受講生からは高評価だった、という話を伺っている。一方で、2科目ともオンデマンド配信授業ということもあり、自身でデータを集め、分析し、活用する機会が少なく、知識として DS や AI の活用方法・事例を知っているだけの状態になっていることが懸念される。たとえば、上記の2科目のうち、春学期開講の「活用法を見聞する AI・データサイエンス」の1回分を筆者が担当している。この科目ではオンデマンド配信による動画を閲覧後、小テストを通して学生の理解度を確認することになっている。筆者が提示した問題のほとんどは知識を問う問題であるが、1問だけ架空のデータを提示してそのデータから言えることを読み取る問題を出したところ、知識を問う問題の平均正答率が約95%だったのに対し、その問題だけ正答率が約60%と、低い数値であった。これは、DS や AI を実際に活用するには知識だけでなく実際にデータを活用する実践的な教育

1) 2021年8月4日に公表された、リテラシーレベルを対象とした教育プログラムの件数である。

2) 2021年度の2科目は「活用法を見聞する AI・データサイエンス」「活用法を体験する AI・データサイエンス」であり、2022年度に追加された2科目は「社会のためのデータサイエンス実践基礎」「AI・データエンジニアリング実践基礎」である。

が必要であることの証左であろう。そのため、全学共通カリキュラムであれば、2022年度から開講される、DSやAIにおける基礎的な実践能力を養う2科目、もしくは各学部で個別に提供されている実践教育プログラム・カリキュラムを受講する必要があると考えられる。

本稿では、大学生が社会の中で使われているAIを使いこなすようになるための教育を実現する一つの試みとして、筆者が所属する社会安全学部で2020年度から提供する実習科目「AI実習」の内容について紹介する。AI実習では、AIには何ができて何ができないのかを体験的に学ぶことで、AIを正しく利活用する能力を身に着けることを目標としている。そのため、情報系の学部にあるようなプログラミングスキルや統計処理で利用されるIBM SPSS Statisticsの活用スキルといった、特定のITスキルの獲得は必ずしも目標とはしていない。AIの得手不得手を理解し、履修後にAIを用いたさまざまなシステム・サービスを正しく活用できるようになることを目指して実習を行っている。

なお、AI実習を実施するにあたり、文理融合学部の特徴である多種多様な学生が存在しITリテラシーに大きな差があること、1年次からも履修可能な科目であり学部カリキュラムの多くが未履修であることなどを踏まえて授業計画を立てている。特に、文系学生に対して、教養としてAIに関する知識を身に着けるのではなく、AIやデータを利活用する力を身に着けてもらうため、理系によくあるプログラミングを軸とした実習とは異なる方式で実習を行っている。ただ、授業評価アンケートからは難しいと感じる履修者が一定数いることがわかっており、まだ多くの観点で工夫すべき点はあると考えるものの、今後のAI教育の参考になればと考え、本稿を執筆する。

2. 社会安全学部におけるIT・AI・DS教育とAI実習の位置づけ

筆者が所属する社会安全学部では、「防災・減災対策や事故防止、危機管理のための政策立案とその実践ができる、高度な情報処理能力を有した社会貢献型人材」^[4]の育成を目標としている。特に、高度な情報処理能力を身に着けるために、本学部では数理やIT技術、統計処理に関する講義・実習を数多く提供している。さらに2020年度からはAI実習をあらたに提供するなど、「AI×社会安全」をキーワードに4年間でAIやDSを学ぶことができる環境を整えている。

AI実習は、配当年次が1年次の科目であるものの、AIを利活用できる能力を身に着けるための実践的な実習科目である。本学部でのAI・DS教育におけるAI実習の位置づけを明確にするために、社会安全学部の学部案内パンフレット（学部案内2022）を参考に、AIやDSに関する一部科目の位置づけを図1に示す。

図1に示す通り、本学部におけるAI・DS教育はリテラシー教育から応用基礎教育までを考えている。つまり、AIを作り出すようなエキスパートな人材というより、AIを用いて社会の課題を解決し、安全・安心な社会を構築することができる人材を育成することを目標と

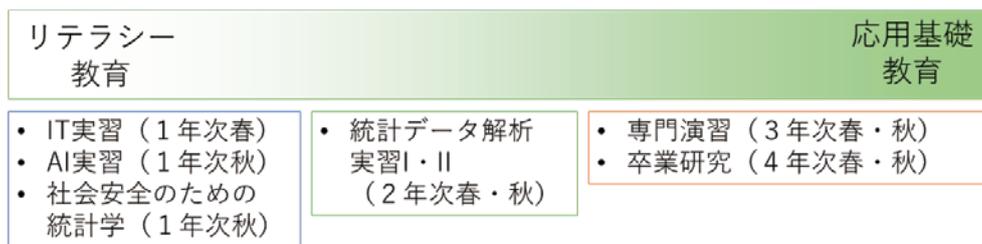


図1 社会安全学部における AI・DS 教育 (一部科目のみ)³⁾

している。そのため、AI 実習は、数理・データサイエンス教育強化拠点コンソーシアム^[5]で述べられている「数理・データサイエンス・AI (リテラシーレベル) モデルカリキュラム」の学修目標⁴⁾、その後半部分である“人間中心の適切な判断ができ、不安なく自らの意志で AI 等の恩恵を享受し、これらを説明し、活用できるようになる”という目標、その先にある「数理・データサイエンス・AI (応用基礎レベル) モデルカリキュラム」の学修目標⁵⁾を達成するための科目となっている。履修者には、3年次・4年次配当である専門演習・卒業研究において、AI 実習や統計データ解析実習 I・IIなどを学習して得られた AI・DS に関する知見をもとに、高度な研究を遂行することが期待されている。

3. AI 活用リテラシーの向上に向けた実習科目「AI 実習」の概要

AI 実習の内容を紹介する前に、AI 実習を履修する学生の特徴を述べ、授業計画を立てる上で考慮すべき事項をまとめる。

まず、AI 実習の履修者の観点から考慮すべき点をあげる。AI 実習は1年次秋学期開講の科目ということもあり、履修者の多くは1年次生である。そのため、1年次春学期に必修科目である IT 実習を通して IT スキル、具体的には PC 操作スキルや Microsoft Office の利用スキルを身に着けたあと、AI 実習を履修することになる。他に1年次春学期の必修科目として、大学での学び方を知る入門演習や安全に関する各種問題を鳥瞰的に学ぶ社会安全学総論 I、語学 (英語科目) があり、選択必修科目として数学の基礎理論を学ぶ社会安全のための数学 I があるが、これらの科目は AI リテラシーとは直接関わりのない科目である。本学が1年

3) 図1では、AI 実習も含めた1年次配当科目がリテラシー教育に相当するような配置になっているが、年次配当で分けているだけである。そのため、モデルカリキュラムの観点からみると、AI 実習は、リテラシーレベルの「選択 (オプション)」の内容を提供しつつ、より発展的な内容である「応用基礎」に関する内容を扱っているため、注意していただきたい。

4) 今後のデジタル社会において、数理・データサイエンス・AI を日常の生活、仕事等の場で使いこなすことができる基礎的素養を主体的に身に付けること。そして、学修した数理・データサイエンス・AI に関する知識・技能をもとに、これらを扱う際には、人間中心の適切な判断ができ、不安なく自らの意志で AI 等の恩恵を享受し、これらを説明し、活用できるようになること^[6]。

5) 数理・データサイエンス・AI 教育 (リテラシーレベル) の教育を補完的・発展的に学び、データから意味を抽出し、現場にフィードバックする能力、AI を活用し課題解決につなげる基礎能力を修得すること。そして、自らの専門分野に数理・データサイエンス・AI を応用するための大局的な視点を獲得すること^[7]。

次春学期に提供する「活用法を見聞する AI・データサイエンス」科目も含め、他の科目はすべて選択科目であり、どの科目を履修するかは学生に委ねられている。以上を踏まえると、AI 実習を履修する学生は一定の PC 操作スキル・Office ソフトの操作スキルがあるだけで、プログラミングや AI に関する知識はないということを想定する必要がある。さらに、第 1 節で述べた通り、本学部は文理融合学部であり、数学の得手不得手、IT スキルの得手不得手が学生間にあることも考慮しなくてはならない。

履修者の特徴以外の観点からは、本授業は 1 単位の実習科目である点や本学が BYOD を推進している点を考慮する必要がある。特に前者は、45 時間（授業時間の学修で 30 時間と授業時間外の学修で 15 時間）の学修時間で単位が取得できるように授業を設計する必要があることを示している。

以上をまとめると、

- ・ Office ソフトを一定程度活用することができる
- ・ IT スキルや数学に関する知識は一定程度あるものの、学生間で差がある
- ・ プログラミングを習っていない
- ・ AI に関する知識を持っていない
- ・ 単位取得のための学修時間は 45 時間である
- ・ 可能な限り関西大学が推奨する BYOD を取り入れる

という点を考慮しつつ、“人間中心の適切な判断ができ、不安なく自らの意志で AI 等の恩恵を享受し、これらを説明し、活用できるようになる”、“データから意味を抽出し、現場にフィードバックする能力、AI を活用し課題解決につなげる基礎能力を修得する”ことを目指して授業計画を立てる必要がある。

3.1 AI 実習の授業概要と実習にあたっての考え方

2022 年度の AI 実習のシラバスに記載されている授業概要の一部を以下に記す。

本実習では、現代の AI の中核技術である機械学習を実装し試すことで、AI は何ができて何ができないのか、AI の本質はどこにあるのかを理解し、AI リテラシーを高めて AI を使いこなすことを目的とする。最初に AI やプログラミングの基礎を説明した後、機械学習の各モデルを実装し、様々なデータに対して実行してその結果を検証する。例題となるデータには、表形式で表される一般的なデータから画像・文字といった各種メディアのデータを取り扱う。その後、実際に各自で用意したデータを用いて自身で機械学習を実行し、結果を検証する。

AI 実習では、現代の AI の中核をなす機械学習、特に教師あり学習に絞って実習を進めている。現代のいたるところで利用されている AI を履修者が使いこなすようになるのが目的のため、画像データや文字データといった、現代の AI を利用した各種サービスで利用されるさまざまなデータを入力し、どのような出力結果になるのか、なぜそのような結果になるの

かを考えさせるようにしている。

このような実習内容にしている理由として、筆者は、AIが搭載された各種システム・サービスの利用者（つまり現代の人々）の、AIに対するブラックボックス化⁶⁾を解消することこそが、AIを正しく使いこなす必須条件であると考えているためである。たとえば、大学生がよく利用するシステムの一つである英語の自動翻訳システムで考えると、一昔前までは日本語をシステムに入力したら明らかにおかしい英語の文章が出てくることが多かったが、現在では深層学習の導入により高精度な英語の文章が出力されるようになっている。そのため、AIの性質を理解し、翻訳システムを正しく使いこなすことができれば非常に有用ではあるが、一方で、日本語を入力したらふさわしくない英語の文章が自動で出力されるため、十分なテラシーがないままでも使えてしまっている。その結果、何も考えずシステムを妄信し、誤訳だったりふさわしくない英語の文章でなかったりしても出力された英語の文章をそのまま利用してしまっているケースが後を絶たない。

現代のAI、特に教師あり学習を基本としたAIについては、人間の学習プロセス・判断プロセスと同じように、教師データをもとに事前に正解・不正解を学習したうえで、未知のデータに対しては学習成果に基づいて結果を判断している。このことだけでも知っていれば、学習したデータ以外のことは正しい判断ができないことや学習したデータにノイズがあれば正しく学習できないこと、正しく学習していても入力時にノイズがあれば正しい結果が返ってこないことなど、AIを利活用するうえでの注意点に気づくことができる。また、そのようなAIだからこそ、人間が求めている結果を出力しているとは限らないこともわかる。この点まで理解できれば、今後AIを搭載した各種サービスを利用した際に結果を妄信せず、利用者自身で結果を正しく判断するようになることが期待される。こうした考えを踏まえ、本実習では、AIの入出力に焦点をあて、履修者に結果の違いを考えさせることにより、AIが何をしているかを理解させ、AIのブラックボックス化の解消を試みている。

3.2 AI実習の授業計画

2022年度のAI実習の全15回の授業計画は次のとおりである。

- 第1回 ガイダンス、AIの歴史、機械学習の適用先と事例紹介
- 第2回 機械学習を行う前の事前準備（1）：Pythonによるプログラミング
- 第3回 機械学習を行う前の事前準備（2）：ライブラリの利用方法とデータの読み込み・可視化
- 第4回 機械学習を学ぶ～機械学習の基礎的なアルゴリズムを学ぼう～（1）：線形回帰、正則化
- 第5回 機械学習を学ぶ～機械学習の基礎的なアルゴリズムを学ぼう～（2）：ロジスティック

6) ここでいうAIに対するブラックボックス化というのは、深層学習のようなモデルが複雑すぎて説明できないことを指しているのではなく、システムへの無理解としたほうがニュアンスは近いかもしれない。

ク回帰

- 第6回 機械学習を学ぶ～機械学習の評価方法と過学習を学ぼう～
- 第7回 機械学習を試す～機械学習の代表的なアルゴリズムを試そう～（1）：サポートベクトルマシン（線形、カーネル法）
- 第8回 機械学習を試す～機械学習の代表的なアルゴリズムを試そう～（2）：テキストデータの扱いとナイーブベイズ
- 第9回 機械学習を試す～機械学習の代表的なアルゴリズムを試そう～（3）：画像データの扱いとニューラルネットワーク
- 第10回 機械学習を試す～機械学習の代表的なアルゴリズムを試そう～（4）：ランダムフォレスト、kNN
- 第11回 機械学習を使う～自分でAIを作ろう～：AIの成功例・失敗例を学ぶ
- 第12回 機械学習を使う～自分でAIを作ろう～：学習のための画像データを収集し加工する
- 第13回 機械学習を使う～自分でAIを作ろう～：様々な機械学習アルゴリズムに適用して結果を考察する
- 第14回 機械学習を使う～自分でAIを作ろう～：失敗原因を取り除いてよりよいAIを作る
- 第15回 総括、課題に対する総評、深層学習に向けて

第1回～第3回までがAIの簡単な歴史やPythonという言語、各種ライブラリの紹介などの、機械学習を実行するまでの準備段階、第4回～第6回がAIの中核をなす機械学習の一つである教師あり学習の基本的な内容、第7回～第10回がさまざまな教師あり学習のアルゴリズムの紹介となっている。また、第7回～第10回の間で実際のテキストデータや画像データの扱いも学ぶようにしている。そのため、第10回までの内容を学習することで、履修者は教師あり学習が行っていることや教師あり学習のアルゴリズムの違い、結果の評価やその見方を理解するだけでなく、自身で身近にあるさまざまなデータに対して教師あり学習を試すことができるようになる。

その後、第11回～第15回の5回をかけて、第10回までのライブラリ付属の実験用データセットや教員側で用意したデータセットなどを使ったわかりやすいサンプルデータを使うのではなく、普段われわれがインターネットを利用する中でよく見るデータ、具体的には画像データを題材に、2種類の画像を判別するAIを一から作ることに挑戦する。第11回時にAIプロジェクトの成功・失敗ポイント^[8]を学習した後、履修者は、画像収集から画像分類AIの完成までのプロセスを通して、サンプルデータや機械学習のサンプルプログラムを試すだけではわからない、AIを活用するために必要なポイントを体験的に学ぶ。これにより、教師データの準備の難しさや“Garbage in, garbage out”に代表されるさまざまなデータ分析・解釈・活用の罫があること、試行錯誤を繰り返し次の学習につなげることの重要性を実感できるようにしている。

3.3 AI 実習の授業形態と主な利用ツール・書籍

コロナ禍であるものの、AI 実習は、高槻ミューズキャンパスにある IT 教室（PC の OS は Windows、教室利用時の様子は図2）を利用して対面で実施している。そのため、PC を持っていない学生でも受講可能であるが、関西大学では BYOD（Bring Your Own Device）を推奨していることもあり、自身の PC でも試すことができるように授業設計している。そこで、授業設計当初は、MATLAB などほかの環境の利用も考慮に入れたものの、OS に依存することなく無料で利用可能な Anaconda を利用することにした。

Anaconda はデータ分析や機械学習など、DS や科学技術計算を行うための各種ライブラリやツールがまとめられている Python 言語のディストリビューションである。世界中で広く使われていることもあり、情報源も豊富にある。たとえば、Anaconda のインストールであれば、指定した教科書にも OS ごとに記載があり、インターネットで検索すれば簡単に詳細な情報を見つけることができる。そのため、特定の環境（ここでは IT 教室の PC）に依存せず、それぞれの学生環境に合わせて進めることができる。

また、授業を効率よく進めるため、e-Learning ツールの一つである関大 LMS を使っている。一般的な e-Learning ツールと同じく、教材の提示・公開やレポート提出、小テスト、ミニッツペーパー、出欠管理などができるようになっている。AI 実習では、各種資料の公開、課題の提示、レポート提出、出欠管理を本ツールで行っている。なお、ミニッツペーパーのみ、Google フォームを用いて実施している。

本実習では、教科書として「見て試してわかる機械学習アルゴリズムの仕組み 機械学習図鑑」^[9] を利用している。機械学習のアルゴリズムや評価法、各種データの取り扱いなど、幅広い分野を網羅していること、カラーでの図解が多くあり直感的にわかりやすく Python のソースコードもシンプルであることなど、機械学習の全体像を俯瞰しやすい内容となっている。一方で、直観性を重視しつつ広く浅く学ぶ構成になっていることからアルゴリズムに対する数学的な説明はほぼなく、より深く理解するにはほかの書籍やサイトを参考にする必要があること、シンプルなプログラムゆえ扱っているデータが一部簡素化されており現実のデータとは異なること、プログラミングの経験がある程度ある人を対象としていることから Python 言語や利用ライブラリの説明は最低限であることなど、本書籍をなぞるだけでは AI 実習の履修者にとってはハードルが高いものとなっている。前半の理解不足に陥るかもしれない箇所に関する補足は教員側で事前に準備することで対応できるものの、後半のプログラミングに関しては、プログラミングを習っていないという履修者の前提条件と相反することから、対応には何らかの工夫が必要である。



図2 IT 教室利用時の様子
(出典：関西大学社会安全学部 HP https://www.kansai-u.ac.jp/Fc_ss/life/campus.html)

3.4 ソースコードを実行する環境と利用ライブラリ

機械学習のプログラムが書かれたソースコードを実行するにあたり、学生自身が所有する PC でも実行可能にするため、Anaconda のインストールで利用可能なものだけを扱うようにしている。そこで、ブラウザ上で Python などのプログラミング言語を記述・実行可能な Jupyter Notebook を利用している。拡張子が「ipynb」という、学生にとっては見慣れないファイル形式であるが、図3に示す通り、コンパイルすることなくセル（入力部分）ごとに実行可能なことから1つずつコードを確かめながら作業が進められること、プログラムの記述だけでなく図表も含めた可視化された実行結果も保存できるなどといった特徴がある。そのため、プログラミング初心者である学生にとって扱いやすく、教員にとっても事前に動かし結果の提示やレポート提出時の結果の確認が簡単にできるといったメリットがある。

また、利用するライブラリも、Anaconda のインストールで利用可能なものだけにしている。本実習で利用する主なライブラリは以下の5種類である。

- ・ Scikit-learn：機械学習ライブラリ
- ・ NumPy：多次元配列を扱うライブラリ
- ・ Pandas：データ解析用ライブラリ
- ・ Matplotlib：データの可視化用ライブラリ
- ・ Seaborn：データの可視化用ライブラリ

AI 実習ではテキストデータや画像データなどのマルチメディアデータも対象としている。ただし、日本語のテキストデータであれば形態素解析、画像データであれば特徴量抽出など

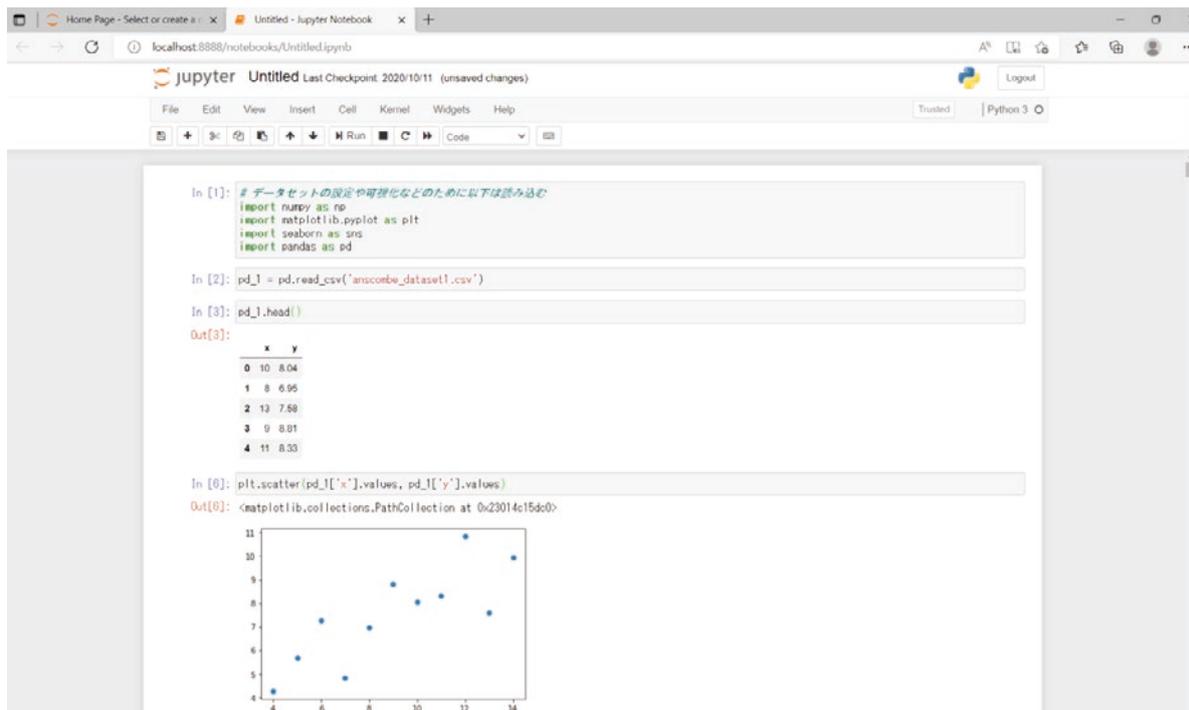


図 3 Jupyter Notebook を用いたプログラミングの実施例。コードだけでなくグラフも含めた実行結果が保存可能。コードは各セルに入力されるため、そのコードだけのやり直しも可能。

の処理は別のライブラリ⁷⁾が必要となることから、授業内では紹介にとどめてこれらの処理は行わず、単純に読み込みした各データをそのまま機械学習に利用するようにしている。そこで、テキストデータに関しては、英語に限定することでデータセット（たとえば英語のスパムメールのデータセット等）をそのまま読み込み、Scikit-learn を用いて機械学習できるようにしている⁸⁾。画像データに関しては、画像の各ピクセル値をそのまま数値として利用、つまりベクトル化させて利用する最も単純な方法をとっている。たとえば、 8×8 のグレースケール画像であれば、64次元のベクトルとなり、この画像が1,797枚あれば、 $1,797 \times 64$ の行列として読み込むことになる⁹⁾。その後、Scikit-learn を用いて機械学習させることになる。

3.5 AI 実習の進め方と授業運営にあたっての工夫

授業の流れは、1) 教員による前回のふりかえり、2) 教員による内容の説明、3) 説明を聞きながら履修者による各種作業（ソースコードの実行など）の実施、4) 授業内で示した

7) 形態素解析であれば mecab、特徴量抽出であれば OpenCV を利用することが多く、どちらも別途インストールが必要である。

8) 自然言語を簡潔に処理する場合、1) 形態素解析、2) 単語のベクトル化（単純な方法として Bag of Words : BoW がある）、という手順になる。そのため、空白で単語が分かれる英語の文章の場合、形態素解析を省略することができる。なお、ベクトル化は必要となるが、BoW は Scikit-learn ライブラリで可能である。

9) これは、Scikit-learn 付属のサンプルデータセットの一つである MNIST の情報である。そのため、MNIST を利用する場合は、画像の読み込みに必要となる Scikit-image なども不要となり、他のサンプルデータセットと全く同じ要領で画像に対して機械学習を実施することができる。

課題の実施、5) ミニッツペーパーの提出、となる。学生側の机の上には、図2の通り、教員の画面を映すモニターが常設されており、履修者はモニターを通して教員の説明を見聴きしながら作業をこなすことになる。また、手を動かす作業が多いため、各種作業の補佐としてTAおよびSAが各一名配置されている。

教員が説明に利用する資料は、関大LMS上に事前に公開している。参考までに、関大LMSにて公開している授業資料の一部を図4、図5、図6にそれぞれ示す。前段落の授業の流れと合わせて説明すると、最初に該当回での目的を示したうえで(図4)、教科書やサイトなどを参考に必要な説明を行い(図5)、サンプルのソースコードを実施して該当モデルで行っていることを確認し、データや各種パラメータを変えるとといった課題を行ってより深い理解を促す(図6)、といった内容となる。図5ではサイトのURLのみを掲載しているが、実際に説明する際は、電子書籍版の教科書をモニター上に提示して説明しており、サイトはその補足説明のために利用している。

成績評価に関しては、3つのレポート課題(小レポート2つと大レポート1つ)を提示し、提出されたレポートの完成度合いで評価している。小レポートは、第10回までの内容に対するレポートであり、授業外で行うものである。一方、大レポートは第11回以降の一から画像分類AIを作る内容に関するレポートであり、授業を通して作成するものとなっている。どの課題も、結果を正しく読み取ったうえで、なぜそのような結果が出力されるのかを考察し、AIは何をしているのかを問う内容となっている。

なお、本節の最初で述べた通り、学生にはプログラミングスキルやAIに関する知識はなく、また、基礎的なITリテラシーにも差があることから、理系学部で行われるプログラミング実習のように進めることは難しく、AI実習の目的から考えてもプログラミングに特化するわけにはいかない。一方で、実習で利用する教科書でもある程度のプログラミング経験がある人を対象としているように、Anaconda(つまりPython言語)を利用して機械学習を実施するためには一定のプログラミングスキルが求められる。さらに、仮にプログラミングができたとしても、多くの機械学習の書籍やサイトに書かれているような、学生の身近にない

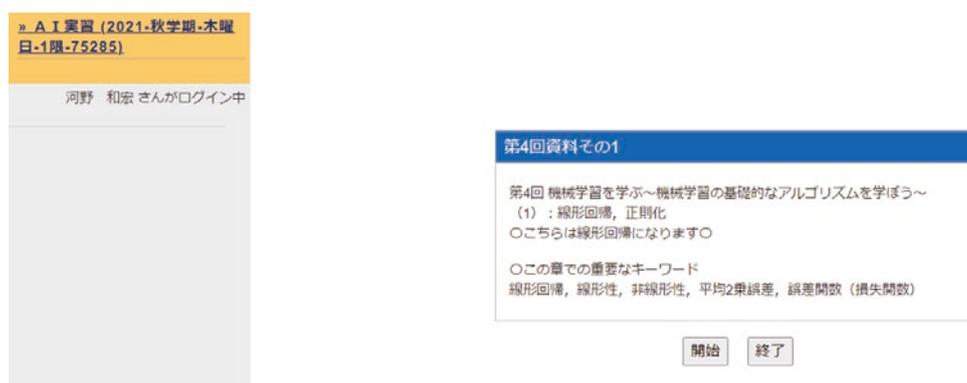


図4 授業資料の一例(1): 第4回における目的とキーワードの提示



図 5 授業資料の一例 (2) : 第 4 回の線形回帰の内容のページ

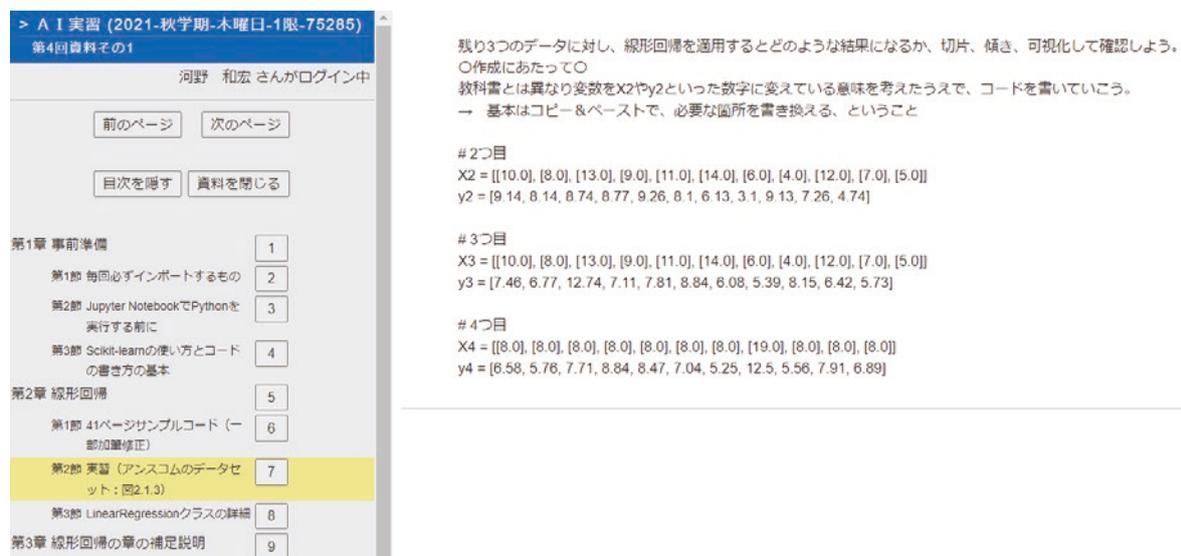


図 6 授業資料の一例 (3) : 第 4 回の線形回帰に関する課題

データばかりを利用し、かつ結果も数値のみでしか出力しないとすると、AI に対する履修者の興味を大きく損なってしまう可能性が高い。そこで、プログラミングスキルがない状態でも滞りなく AI を試すことができ、かつ履修者に AI が身のまわりで使われていることを実感させるために、以下の 2 点を行っている。

3.5.1 ソースコードのパーツ化

AI 実習では、AI は何ができて何ができないのか、結果をどう読み取り活用していくのかに焦点を当てており、AI そのものを構築するための機械学習のプログラミングの習得は目標としていない。そこで、「パーツを組み合わせる」ような感覚で実習が進むように設計している。

同様の考えのものとして、Scratch に代表されるビジュアルプログラミングツールが有名

である。ビジュアルプログラミングツールは、一般的なプログラミングで想像されるテキストコードを入力する形態ではなく、さまざまな動作が書かれたブロックを視覚的に組み合わせる形態であり、正しい文章を作る感覚でプログラミングすることができる。そのため、プログラミング初心者が楽しく学ぶ効果が期待される。AI実習では、Anaconda を利用する関係上、ビジュアルプログラミングツールのような形態では進められないが、事前に各動作をするためのコードを関大LMSに公開し、それらを組み合わせて実行する形態にすることで、自らプログラミングすることなく、実習が進むように構築している¹⁰⁾。図4～図6に例示した第4回時の授業時において、学生に提示したコードを図7に、その実行結果を図8にそれぞれ示す。また、組み合わせて使うことが理解できたであろう授業後半において示している各コードのパーツの一例を図9に示す。

Scikit-learn を用いて機械学習を実施する場合、(1) 必要なライブラリのインポート、(2) 学習データとテストデータの用意、(3) 機械学習のアルゴリズムの指定と学習の実行、(4) テストデータでの検証、(5) 結果の表示、の手順で実行することになる。そのため、機械学習のプログラムは5つのコードに分割することができる。図7で示した第4回時点ではコードを一纏めにして示しているが、まずはこの手順ですべて実施可能であることを意識づけるようにしている。そのため、実行例である図8では、一つずつセルにコードを入力（コピー&ペースト）して示している。また、このような形で見せることで、機械学習そのもののコードは非常に単純であり、むしろ事前準備（データの準備・入力）や事後処理（結果の出力）のコードの方が長くなるということも説明している。加えて、図6のような学習データとテストデータのみを変える課題を提示している理由の一つに、(2)に該当するコードを変更しさえすれば他のコードは同一のコードで実行可能であることを体験させるという意図もある¹¹⁾。

(1)～(5)の各段階でのコードを組み合わせて機械学習ができることを複数回の授業を通して理解させたいと、授業後半やレポートでは、図9に示したコード（＝ソースコードのパーツ）の一覧を関大LMS上で提示し、履修者自身で組み合わせて機械学習を実施するようにしている。たとえばランダムフォレストやkNNを実施する第10回では、1) 必要なライブラリのインポート（図9における第1章に該当し、第1節のコードを利用）、2) 学習データの用意（図9における第2章に該当し、データ入力・作成のパターンに応じてコードを利用）とテストデータの用意（図9における第3章に該当し、第1節のコードを利用）、3) および4) 今回利用する機械学習のアルゴリズムであるランダムフォレストとkNNの実行とテストデータでの検証（図9における第5章に該当し、第8節と第9節のコードを利用）、5) 結果の表示（図9における第6章に該当し、各節から必要なコードを利用）、といった順に、関大

10) 近年では、GUIベースでAIを構築可能なサービスもいくつか登場しているが、授業の根幹が外部サービスに依存してしまうのは望ましくないと考えられる。

11) 第4回の課題の場合、厳密には変数名やモデル名を変える必要があるため、全く同じコードというわけではない。これは、各変数やモデル名がどのコードの部分に対応しているかを確認してもらうために、あえて書き換えるような課題にしている。



図 7 第 4 回で示した線形回帰のコード。機械学習は 5 つの手順で実施されることが示されている。また、(2)や(5)に比べ、機械学習のアルゴリズムの実行部分である(3)や(4)のコードは非常に簡単である

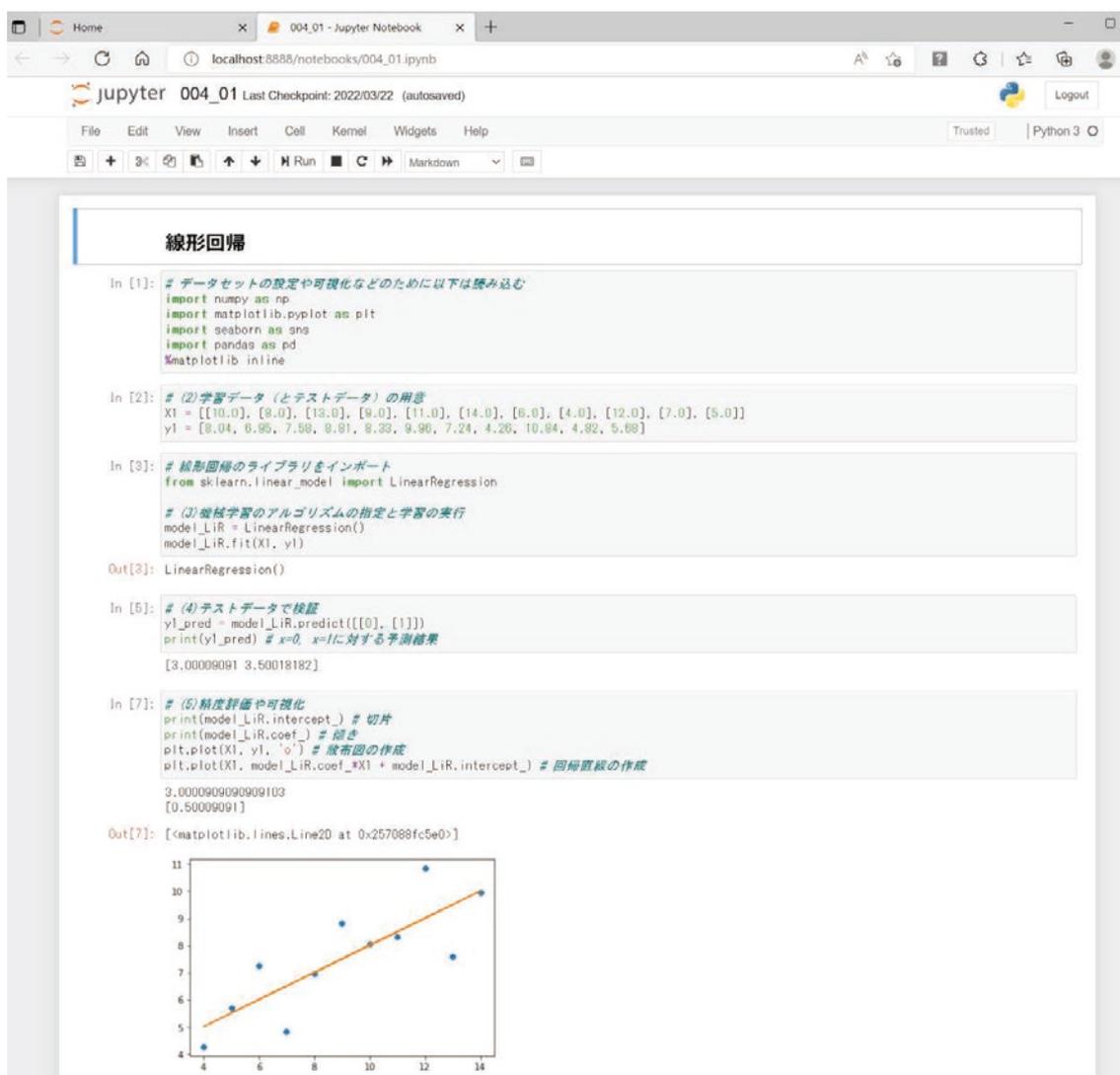


図 8 線形回帰の実行例。関大 LMS 上で公開しているコードをそのままコピー & ペーストすればだれでも実行可能である

LMS の各ページにアクセスし、コードをコピー&ペーストして利用することになる。

なお、各コードを組み合わせて利用する関係上、関大 LMS で示す各節のコードは各章間でつながるよう、各種変数名は統一しているが、機械学習の実施の際のモデル名のみ、それぞれ固有のモデル名にしている。これは、現在何のモデルで実行しているかを明確にするためである。それゆえ、結果の表示の際にモデル名が必要となる場合は「modelclassifier」として記載し、その部分のみ書き換えるようコメントを入れている。こうした書き換えが必須な部分は、関大 LMS が HTML で書かれていることを利用して赤字でわかりやすく表示している。

ここで、学生に提示するコードの多くは非常に簡単なものであるため、一部のコードは一から入力させてもよいのでは考える読者もいるかもしれないが、授業運営の観点からプログラミング未経験者には極力入力させないほうが好ましいと考えている。実際、本学部で過去に担当した IT 活用実習 2 という、統計処理のスキルを身につける 2 年次配当科目において、IBM SPSS Statistic と R を用いた実習を実施したところ、すべて GUI (Graphical User Interface) 形式で行うことができる IBM SPSS Statistic はある程度スムーズに進められたものの、一部の処理にコマンド入力、つまり CUI (Character User Interface) 形式で入力が必要となった R の場合、たった一文のコマンド入力だけでも頻繁に手が挙がり、授業が全くスムーズに進まなかった、という経験がある。プログラミングに慣れていない場合、ちょっとした打ち間違いに気づかない、間違っただけのエラー文が理解できないなど、うまく動かなかった時点で思考を放棄する傾向にあるようである。そのため、今回のようにコードは教員側ですべて用意し、あとは組み合わせるだけの方が、教員の想定外のエラーが出ることなく、授業をスムーズに運営することができると思われる。

> A I 実習 (2021-秋学期-木曜日-1限-75285)
機械学習の実行コード一覧

河野 和宏 さんがログイン中

前のページ 次のページ

目次を隠す 資料を閉じる

第1章 事前準備	1
第1節 よく利用するライブラリのインポート	2
第2節 Jupyter NotebookでPythonを実行する前に	3
第3節 Scikit-learnの使い方とコードの書き方の基本	4
第2章 データの用意	5
第1節 sklearnのデータセットを利用する場合	6
第2節 (参考) sklearnのデータセットをデータフレーム形式に変更したい場合	7
第3節 関数を利用して一から作成する場合	8
第4節 csvファイルを準備して読み込む場合 (パターン1)	9
第5節 csvファイルを準備して読み込む場合 (パターン2)	10
第6節 csvファイルから読み込む場合の注意点	11
第3章 データの分割	12
第1節 学習データとテストデータへの分割	13
第4章 データの中身の確認と可視化 (任意)	14
第1節 matplotlibによる2次元データの可視化	15
第2節 seabornによるデータフレーム形式のデータの可視化	16
第5章 アルゴリズムの指定と学習, テスト	17
第1節 回帰分析 (線形)	18
第2節 リッジ回帰 (正則化)	19
第3節 ロジスティック回帰	20
第4節 SVM (線形)	21
第5節 SVM (カーネル利用: 非線形)	22
第6節 ナイーブベイズ (多項分布モデル)	23
第7節 NN: ニューラルネットワーク	24
第8節 ランダムフォレスト	25
第9節 kNN: k-近傍法	26
第6章 結果の表示(2クラス)	27
第1節 分類の場合その1 (各種精度の出力)	28
第2節 分類の場合その2 (決定境界の可視化)	29
第3節 分類の場合その3 (ROCカーブの出力)	30
第4節 混同行列の意味と正解率、適合率、再現率の関係	31

- アルゴリズムの指定とモデルの学習, テスト
- (0)事前のデータの確認 ※ここは(2)の後になる場合もあります
 - (1)必要なライブラリのインポート ※いつでも可ですが忘れないようにする
 - (2)学習データとテストデータの用意
 - (3)機械学習のアルゴリズムの指定と学習の実行
 - (4)テストデータで検証 ※過学習・未学習の確認のため学習データでも実施
 - (5)精度評価や可視化

○注意

- ・データを分割した前提で進めているため、もし分割せずそのまま全データをトレーニング用を利用する場合は、読み替えること

ロジスティック回帰を選択した場合

```
# ロジスティック回帰のライブラリをインポート
from sklearn.linear_model import LogisticRegression

# モデルの定義と構築
# モデル名は「model_LoR」としている
# LogisticRegressionのパラメータ等は第5回を確認
model_LoR = LogisticRegression()
model_LoR.fit(X_train, y_train)

# テスト
y_pred_train = model_LoR.predict(X_train)
y_pred_test = model_LoR.predict(X_test)
```

図 9 機械学習を実行するコードの一覧。各章から必要なコードが書かれている節を選択し、そのコードをコピー＆ペーストして利用できるようにしている。章を選択すると、この図のように5つの手順のうちの該当箇所が赤字で示され、章内の各節（たとえばロジスティック回帰）を選択すると、四角枠内のようにそのコードが示される

3.5.2 学生の興味関心を促すテーマ設定

AI実習では、学生の興味関心を喚起させるようなテーマを題材に機械学習を試すようにしている。そのため、Scikit-learnなどで用意されているサンプルデータや統計処理で使われるアンケートデータだけでなく、学生がよく見聞きする内容に関するデータ、身近にあるデータを利用するようにしている。たとえば、小レポートの課題のうちの一つは、ポケットモン

スターにおける各ポケモンのデータ（csv ファイル）¹²⁾を用いた課題となっている。架空のデータを用いるのではなく、多くの学生が知っているであろう内容のデータを利用することで、結果の考察の理解を手助けするだけでなく、身の回りのデータにも簡単に適用でき、かつ、実際に学生の身の回りで AI は利用されているということを実感させるようにしている。

第11回以降の授業や大レポートでは、2種類の画像を判別する AI を作る内容にしている。日常で頻繁に利用するデータである画像データを扱うことで親近感がわきやすく、かつ、出力結果が視覚的にわかりやすい、という効果が期待される。さらに、教員側で画像データを用意するのではなく、学生が興味ある2種類の画像データに対して機械学習を実行させることで、より興味関心を喚起させるようにしている。これまでの履修者が提出したレポートの内容を見ても、2種類の動物を判別するといったありふれた AI から、アニメの中での登場人物二人を判別する AI、ゆるキャラのイラストとそのキャラの元となる人物・動物の画像を判別する AI、兄弟や姉妹を判別する AI、2種類の風景（雨と雪など）を判別する AI など、履修者が判別してみたい画像に対して AI を作り上げようとしている。

3.6 2種類の画像を判別する AI の作成方法と留意事項

第3節の最後に、第11回以降の授業の内容および大レポートの内容である、2種類の画像を判別する AI を一から作る手順および留意事項を紹介する。

3.6.1 画像の収集および加工

AI を構築する前に、学習用・テスト用のデータを準備する必要がある。そこで、1) 画像のダウンロード、2) 画像の jpeg 変換、3) 画像のリネーム、4) 画像のサイズ変更、を第12回で実施し、画像データの収集およびフォーマットの統一を行っている。これらは Python のプログラムではなく、さまざまなフリーソフト・アドオンを利用することで実現している。また、これらはすべて IT 教室の PC だけでなく、自身が所有する PC でも利用できるようにするため、可能な限り OS（Windows と Mac）問わず実行できるようにしている。

各段階での注意点を述べると、1) 画像のダウンロードでは、Google Chrome などのブラウザ上で動作する、画像を一括ダウンロードするアドオンを利用している。この理由の一つは、解放ポートの関係により大学内のネットワークでは専用ソフトでダウンロードできない場合があるためである。また、ダウンロード画像の枚数は1種類あたりだいたい500枚～600枚ぐらいになるが、本実習で利用したアドオンの場合、画像と全く関係のないファイル（ブラウザのアイコンや svg などの通常の画像フォーマットとは異なるファイル形式のものなど）がダウンロードされるケースがあり、これらはノイズにしかならないのでこの段階で削除するようにしている。

12) ポケモンのデータは、Kaggle や GitHub などの各種サイトで公開されている。

次に、2) 画像の jpeg 変換¹³⁾、3) 画像のリネーム¹⁴⁾、4) 画像のサイズ変更は、収集した画像をある規則に従って命名し、かつ統一したフォーマットにするためにフリーソフトを利用して実行している。収集した画像は、ほとんどが jpeg 画像ではあるものの、一部に png 画像もあるため、jpeg 形式の24bit 画像にすべて変換している。その後、AI 構築時にループ処理で画像を読み込ませるために、各画像の名前をリネームし、ラベルと番号で名付けている。たとえば、ラベル1の1つ目の画像なら、L01_001という名前に変更している。最後に、画像のサイズをすべて統一したサイズに変更する。授業中では、64×64 (ピクセル) の画像になるように変更している。2種類の画像を利用するため、ここまでの作業を2回行うことになる。

3.6.2 画像を判別する AI の作成(1): アルゴリズムの決定

第13回では、画像を判別する AI を試作する。評価指標に対しての達成基準 (たとえば正解率が90%以上) を決め、図9に示したコードの一覧を使って構築する。ただし、画像の入出力に関しては、図9のコードでは対応できないため、教員側で別途用意・配布している¹⁵⁾。つまり、(2) 学習データとテストデータの用意に関しては教員側が関大LMSを通じてコードを別途配布し、それを実行した後、これまで通り (3) 機械学習のアルゴリズムの指定と学習の実行以降の処理を行うことになる。その後、(5) 結果の表示において各評価指標を出力後、関大LMSで別途配布した画像分類の結果を出力するコードを実行することになる。

ある機械学習を1回実行した際のフォルダの構成は図10のようになる。Original フォルダ内には入力画像のフォルダがあり、train や test のフォルダは学習およびテストに使った画像のフォルダ、result_train や result_test のフォルダは学習データおよびテストデータに対して機械学習を行った際の出力結果を指す。result_txt は、機械学習を行った際のラベルの振り分けの結果である。result_test の01フォルダ¹⁶⁾および result.log.txt の一例を図11に示す。result.log.txt 内の左列は機械学習した結果を示しており、L01_002.jpg は正しくラベル1として判別しているが、L01_007.jpg はラベル2として誤って判別されていることがわかる。また、result_test 内の01フォルダを見ることで、どの画像がラベル1として判別されたか一覧することもできる。

なお、ここまでは、ある機械学習のアルゴリズムを1回実施しただけである。そのため、result_train や result_test、result_log.txt が上書きされないよう、別のフォルダに移動させ

13) もともと jpeg 形式である画像も含め、全ての画像を、ソフトを使って一括変換している。なお、グレースケール画像も24bit 画像になるよう色数を設定しておかないと、グレースケール画像で処理されてしまい、読み込み時にエラーとなるので注意が必要である。

14) 履修者のPCによっては、デスクトップなどの主要なフォルダがOneDrive下にあるが、ソフトによってはフリーズの原因になるので注意が必要である。

15) 画像処理ではOpenCVが有名ではあるが、別途インストールが必要となるため、ここではPillowやScikit-imageを使って実現している。両方とも、Anacondaをインストールすれば利用可能である。

16) 出力プログラムの関係で、フォルダ名は「1」となっている

```

Desktop -> pict -> original -> 01 -> 各種画像
                                     -> 02 -> 各種画像
      -> train   -> 01 -> 各種画像
                                     -> 02 -> 各種画像
      -> test    -> 01 -> 各種画像
                                     -> 02 -> 各種画像
      -> result_train -> 01 -> 各種画像
                                     -> 02 -> 各種画像
      -> result test -> 01 -> 各種画像
                                     -> 02 -> 各種画像
      -> result.log.txt

```

図10 各画像フォルダの構成例

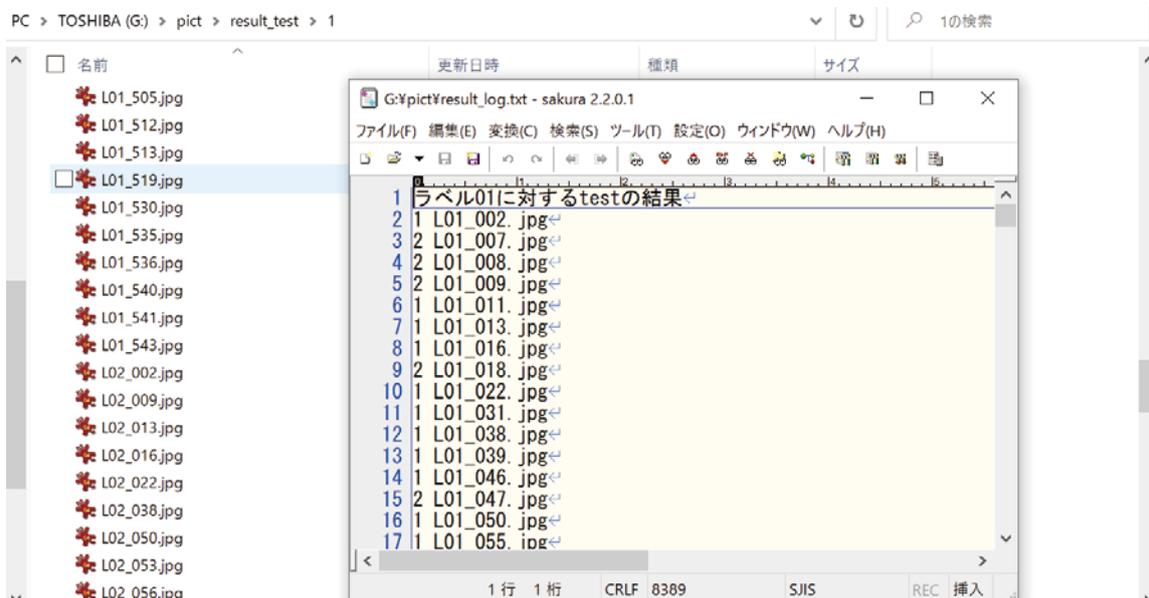


図11 result_test フォルダおよび result_log.txt の中身の一例

たとえば、(3) 機械学習のアルゴリズムの指定と学習の実行以降の処理を、違う機械学習のアルゴリズムで実行する。その後、図12に示す各指標の結果も含めた、各アルゴリズムからの出力結果およびその理由を考え、どのアルゴリズムが画像判別にふさわしいアルゴリズムなのか検討する。

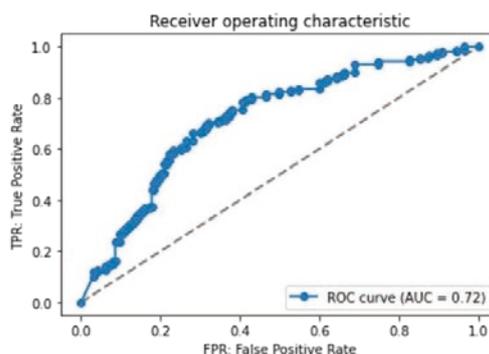
3.6.3 画像を判別する AI の作成(2)：データに含まれるノイズの除去

第14回では、目標値を達成する画像判別 AI の構築を目指す。アルゴリズム決定の段階では、ほとんどの場合、目標値より低い判別しかできない AI となっている。この原因の一つは、収集した画像のほぼすべてを学習データ・テストデータに利用したことによる、データに含まれている多くのノイズデータである。画像収集する際は、単に検索キーワードに合致した画像をそのまま利用しているため、さまざまな種類の画像が入っており、その分多くの

ロジスティック回帰

```

テストデータ時の混同行列
[[112 51]
 [ 49 108]]
学習時の混同行列
[[379 1]
 [ 1 364]]
テストデータ時の正解率: 0.6875
学習データ時の正解率: 0.9973154362416108
テストデータ時の適合率: 0.6956521739130435
学習データ時の適合率: 0.9973684210526316
テストデータ時の再現率: 0.6871165644171779
学習データ時の再現率: 0.9973684210526316
テストデータ時のF値: 0.6913580246913581
学習データ時のF値: 0.9973684210526316
    
```



ランダムフォレスト

```

テストデータ時の混同行列
[[112 51]
 [ 45 112]]
学習時の混同行列
[[379 1]
 [ 1 364]]
テストデータ時の正解率: 0.7
学習データ時の正解率: 0.9973154362416108
テストデータ時の適合率: 0.7133757961783439
学習データ時の適合率: 0.9973684210526316
テストデータ時の再現率: 0.6871165644171779
学習データ時の再現率: 0.9973684210526316
テストデータ時のF値: 0.7000000000000001
学習データ時のF値: 0.9973684210526316
    
```

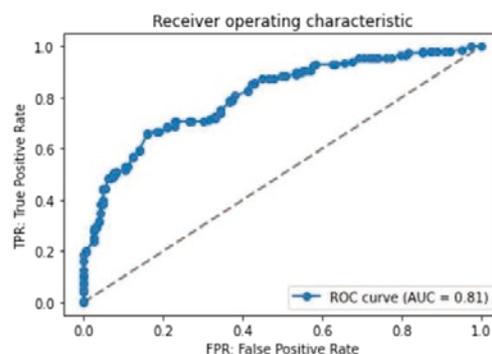


図12 さまざまなアルゴリズムにおける画像判別の結果の一例

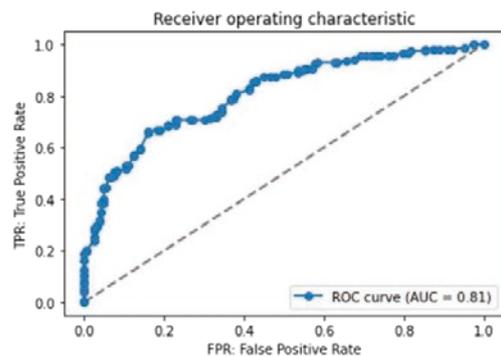
ノイズデータを含んでいる。たとえば人物の画像であれば、理想的な画像はキーワード検索した人物が真正面を向いて一人で大きく写っている画像といえるが、該当人物以外の人物が大きく写っている画像、集合写真のように該当人物が写っていても小さくなっている画像、人物に関係のないモノが多く写りこんでいる画像、人物に関連しているが人物が写っていない全く別の画像、人物とは全く関係がない画像など、人なら総合的に判断できるかもしれないが今回作成する AI にはノイズとなりえる画像が多く存在する¹⁷⁾。それらの画像を、目で見ても original 画像のフォルダから取り除き、もう一度機械学習を実行しなおし、図13に示すように、結果を検証する。その後、まだ目標値を達成していない場合は、データを再び見直してノイズがないのかチェックし、ノイズを再び除去した original 画像フォルダを用いて再び機械学習を実行しなおす。これらを何度も繰り返し、よりよい AI を作りあげていくことになる。

なお、目標値はあくまで履修者が決定した基準であるため、達成・不達成という結果は重要視していない。仮に目標値を達成していても、結果を検証するのをやめるようなことはさせず、収集データ（学習データ・テストデータ）にはまだ問題がないのか、目標値が簡単ではなかったのかなどを考えさせ、よりよい AI を作るように促し、もしそれ以上によりよい

17) キーワード検索による画像収集の関係上、同一の画像が複数存在する場合がある。本来は削除して一つにすべきであるが、実習時間の関係でこれらは削除対象としてない。

前回のランダムフォレストの結果(初回)

```
テストデータ時の混同行列
[[112 51]
 [ 45 112]]
学習時の混同行列
[[379  1]
 [ 1 364]]
テストデータ時の正解率: 0.7
学習データ時の正解率: 0.9973154362416108
テストデータ時の適合率: 0.7133757961783439
学習データ時の適合率: 0.9973684210526316
テストデータ時の再現率: 0.6871165644171779
学習データ時の再現率: 0.9973684210526316
テストデータ時のF値: 0.7000000000000001
学習データ時のF値: 0.9973684210526316
```



2回目のランダムフォレストの結果

```
テストデータ時の混同行列
[[116 27]
 [ 38 74]]
学習時の混同行列
[[332  1]
 [ 0 259]]
テストデータ時の正解率: 0.7450980392156863
学習データ時の正解率: 0.9983108108108109
テストデータ時の適合率: 0.7532467532467533
学習データ時の適合率: 1.0
テストデータ時の再現率: 0.8111888111888111
学習データ時の再現率: 0.996996996996997
テストデータ時のF値: 0.7811447811447811
学習データ時のF値: 0.9984962406015038
```

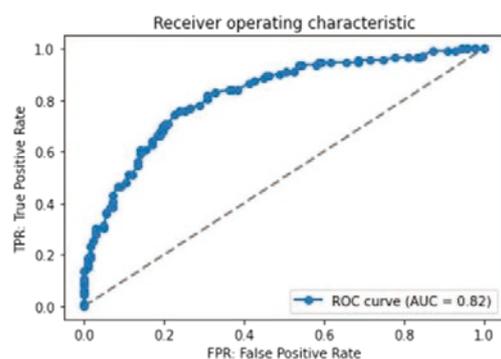


図13 ノイズと思われる画像を除去した際の結果の一例。この作業を何回も繰り返し、AIの性能をあげる。学生のレポートの内容を見ると、だいたい4・5回繰り返している

AIを作るのが難しく性能の限界とを感じるなら、なぜそう感じたのか考えさせるようにしている。これは、最終的に目標値が不達成であっても同様であり、目標値に近づけるように何度も検証を続けた後、目標値を達成するAIを作るのが難しいと感じるならなぜそのように感じたのかを考察させるようにしている。

こうした経験をすることで、AI開発・活用の成否をわける最も重要な要素である、利用するデータの質・量の重要性を認識させ、もしデータにノイズがあれば、“Garbage in, garbage out”という結果につながってしまうこと、そうならないようにするためにも、AIのプログラムそのものよりも、AIに読み込ませるデータの処理(前処理)に十分な時間を割く必要があることを実感させる。加えて、学習データに深く依存してしまうAIだからこそ、必ずしも人がほしい結果を出力しているとは限らず、かつ性能にも限界があり、どのような結果であってもきちんと人の手で検証することの必要性を実感できるようにしている。

4. 今後の課題と授業内容の改良案

本稿執筆時点では、2020年と2021年の2年において、AI実習を実施した。レポート内に記載された感想を見たり、直接履修者に話を聞いたりしたところ、AIが何をしているかやデータがどれだけ重要か、そのためにどれだけ時間をかける必要があるのかといった、教員が

説明した話が実習を通してよく理解できたといった意見、最終的には人の判断が重要であることを実感できたという意見、さまざまなデータを扱えて楽しかったという意見など、好意的な意見がいくつか見られた。

その一方、想定していた通り、授業評価アンケートからは一部の履修者から難しいという意見も見られた。具体的な分析はしていないものの、前期の Office ソフトの利用を学ぶ IT 実習でも、毎年、履修者の60%前後が難しいと感じているようで、難しさの原因の根本には履修者の IT リテラシーの低さ、特に表計算ソフトに代表される数値の取り扱いに対するリテラシーの低さがあると考えている。AI 実習は IT 実習と異なり選択科目のため、AI に興味がある履修者であるものの、すべて教員側でコードを用意しているとはいえ、プログラミングやそこから出力される結果(数値)を読み取るという作業は、一部の履修者には難しかった模様である。

2022年度より高等学校において「情報 I」が共通必修科目になり、2025年度入学生からはある程度 IT リテラシーが身につけている学生が増加すると考えられるものの、しばらくはこの状況が続くと予想される。一部の履修者が難しいと感じることをなくすことは難しいと考えるものの、そのような履修者であっても最後までしっかりと内容を理解させ、筆者が伝えたいことを把握してもらうために、2022年度は、過去2年の実施方法を踏まえたくて、以下のように改善して実施する予定である¹⁸⁾。

1つ目は、学生に対する説明資料の用意である。どちらの年度にもおいても、説明する際は主に電子書籍をモニター上に提示して説明し、時折サイトを用いて補足説明していたが、履修者にとっては、提示資料を行き来するよりも、教科書とサイト両方をまとめたプレゼン資料を用いて説明した方がわかりやすいと考えられる。そこで、2022年度は、IT 実習と同じく全15回、説明用のパワーポイント資料を準備する予定である。また、資料作成に伴い、「文系 AI 人材になる：統計・プログラム知識は不要」^[10]の内容も参考にする予定である。

2つ目は、履修者に提示している各コード(パーツ)に関する組み合わせのフローチャートやマニュアルの整備である。図9に示したように、現在提示しているコードでも、コメントの挿入などでそれぞれ何をしているかわかるようにしているが、多くの章・節がある関係で一目ではわかりにくくなっていたり、コメントだけでは振り返りが難しくなっていたりする点は否めない。そこで、別途、コード利用のためのフローチャートを用意し、どの手順で実施していくか、その際にどのコードを利用するかを明示する予定である。また、各段階のコードに対する簡単な解説を記したマニュアルも準備し、各授業回の内容まで戻らずとも、基本的な内容はそのマニュアルで理解できるようにする。

3つ目は、ソースコードの完全なパーツ化である。図7にあるように、一部の授業回では一括してソースコードを示したり、教科書に記載されたソースコードの関係で必ずしも組み

18) これらの事項は、本稿執筆時点ですでに完了している。

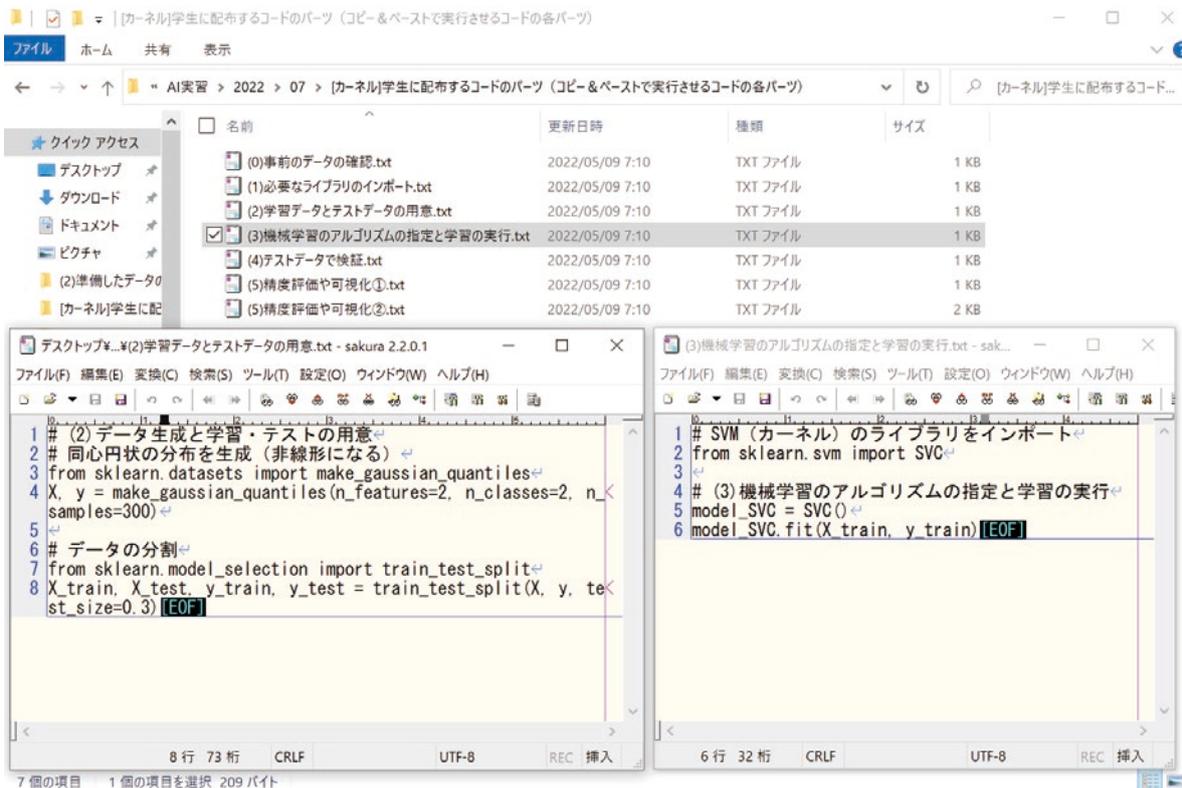


図14 第7回のソースコードのパーツ化と各段階におけるコードの中身

合わせ通りの順番で提示できていなかったりしていた。そこで、全ての回で、フローチャートに沿った手順で必ず実行できるように、各回のソースコードを修正し、かつ、それぞれの段階ごとにソースコードを分割して提示する予定である。一例として、図14を示す。ソースコードはすべて分割され、また、各段階でコードを実行する関係で、該当のコードを実行するために必要なライブラリのインポートは、そのコードの直前に行うように修正している。このような形で学生に提示することにより、より一層、コードを組み合わせて作るという意識づけができると考えている。

5. おわりに

本稿では、筆者が担当する実習科目「AI実習」の詳細について紹介した。本実習では、AIに対する利用者のブラックボックス化の解消を目指し、現代のAI、特にAIの中核をなす機械学習（教師あり学習）は何ができて何ができないのか、AIを利活用するためには何が求められているのかを理解してもらえるように構成している。そのため、プログラミングに焦点を当てるのではなく、出力結果をどう読み取ったのか、そのような出力結果になったのはなぜなのか、よりよいAIを作るためにはどうしたらよいかを考えさせるようにしている。何度も考えさせる・検証させるプロセスを通じ、現代のAIは人が最終的な判断を下す必要があるということ、つまり、AI任せにすることは決してできず、AIを使いこなすには相応の

判断能力がいるということを実感できるような実習内容となっている。

近年はどの大学でも AI やデータサイエンスに関する科目を提供するのは必須となっている。しかしながら、各大学の事情（学生の特徴、教員のレベル、etc.）は大きく異なるため、それぞれ手探りの状態で AI やデータサイエンス教育を進めているということが推察される。AI 実習での工夫として、主に、1) ソースコードのパーツ化、2) 履修者に興味関心を喚起させるテーマ設定、の2つがあるが、今回紹介した内容が AI リテラシー教育の参考になれば幸いである。

参考文献

- [1] 内閣府、“AI 戦略—科学技術・イノベーション—、”
<https://www8.cao.go.jp/cstp/ai/index.html>（2022年3月15日アクセス）。
- [2] 文部科学省、“数理・データサイエンス・AI 教育、”
https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00001.htm（2022年3月15日アクセス）。
- [3] 関西大学、“プレスリリース「学部問わず！2021年4月から全学データサイエンス教育プログラムが始動」を配信、”
<https://www.kansai-u.ac.jp/ja/about/pr/news/2021/01/20214.html>（2022年3月15日アクセス）。
- [4] 関西大学社会安全学部、“設置の理念・目的、”
https://www.kansai-u.ac.jp/Fc_ss/about/philosophy.html（2022年3月15日アクセス）。
- [5] 数理・データサイエンス教育強化拠点コンソーシアム、
<http://www.mi.u-tokyo.ac.jp/consortium/index.html>（2022年3月15日アクセス）。
- [6] 数理・データサイエンス教育強化拠点コンソーシアム、“数理・データサイエンス・AI（リテラシーレベル）モデルカリキュラム～データ思考の涵養～、”
http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf（2022年3月15日アクセス）。
- [7] 数理・データサイエンス教育強化拠点コンソーシアム、“数理・データサイエンス・AI（応用基礎レベル）モデルカリキュラム～AI×データ活用の実践～、”
http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_ouyoukiso.pdf（2022年3月15日アクセス）。
- [8] 増田博之、“失敗から学ぶ機械学習用、”
<https://www.slideshare.net/HiroyukiMasuda/1/ss-181844477>（2022年3月15日アクセス）。
- [9] 秋庭伸也、杉山阿聖、寺田学、加藤公一（監修），“見て試してわかる機械学習アルゴリズムの仕組み 機械学習図鑑、” 翔泳社、2019年。
- [10] 野口竜司、“文系 AI 人材になる：統計・プログラム知識は不要、” 東洋経済新報社、2019年。

