

データセットの準備と機械学習による分類と回帰の実践 -2023 年に新設したデータサイエンス基礎Ⅱの授業開発- Data Set Preparation and Machine Learning Classification and Regression Practices -Development of a new Data Science for Beginners II in 2023-

工藤 知草, 谷口 進一
KUDO Tomoshige, TANIGUCHI Shinichi

金沢工業大学では、2023 年度の前学期から、必修科目として「データサイエンス基礎Ⅱ」の授業を新規に 2 年生対象に開講した。機械学習の教師あり学習としてクラスター分析、教師なし学習として決定木分析やニューラルネットワークなどの手法を用いて分類や回帰を実践する演習をおこなった。さらに、気象庁からダウンロードしたデータから自分でデータセットを準備し、機械学習の教師あり学習により分類と回帰を実践する演習を行った。これは、プロジェクトデザインや卒業研究に機械学習の手法を導入する手助けになる。また、授業動画を充実させた。ここでは、新規に開講した授業の内容とその意義について報告する。

キーワード：データサイエンス, 機械学習, クラスター分析, 決定木分析, ニューラルネットワーク

Kanazawa institute of technology started lectures on Data Science for Beginners II as a new compulsory course for second-year students in the first semester of the 2023 academic year. Exercises were conducted to practise classification and regression using methods such as cluster analysis as supervised learning in machine learning and decision tree analysis and neural networks as unsupervised learning. Furthermore, the students prepared their own datasets from data downloaded from the Japan Meteorological Agency and practised classification and regression using supervised learning in machine learning. This will help to introduce machine learning methods into project design and graduation research. In addition, lecture videos were enhanced. This section reports on the content and significance of the new lectures.

Keywords: Data Science, Machine Learning, Cluster Analysis, Decision Tree Analysis, Neural Networks

1. はじめに

2019 年の AI 戦略¹⁾において、小学生・中学生には基礎的学力・情報活用（100 万人/年）、高校生には数理・データ関連教育（100 万人/年）、大学生には AI リテラシー教育（50 万人/年）を実施することが求められた。文系・理系に関わらず、教養の一つとして AI とデータサイエンスの素養を身につけることが重要な課題となっている。本学では、全学部の必修科目として「データサイエンス基礎Ⅱ」の授業を 2023 年度の後学期から新規に開講した。

ここで、データサイエンスとは、ビッグデータから有益な知見を見出す研究分野で、手法として数学、統計学、機械学習、計算機科学などが用いられる。この授業では、その中の手法の一つである機械学習に注目した。機械学習は、データ内に隠れているルールやパターンを見出し、未知のデータにもルールやパターンに沿って分類と予測を行うモデルを構築する技術である。専門的に機械学習のコードを開発するためには、Python でコードをつくることが望ましく、近年では、GitHub に多くの Python のコードが公開されている。また、データセットも多くある。さらに、Kaggle の Competition に参加して、機械学習の Python のコード開発の能力の研鑽を積むこともできる。

新規に開講した「データサイエンス基礎Ⅱ」は、約 1,600 名の 2 年生全員を対象とし、情報を専門としない学科の学生が多く受講する。これまでにコードを全く開発した経験がない学生が、わずか 100 分×7 回の授業で、ゼロから図 1 で示した機械学習の教師あり学習と教師なし学習の能力を身につけることを可能とした授業を新規に設計する必要があった。本学は、MATLAB Campus-Wide ライセンスを有している。ここで、MATLAB はエンジニアや技術者向けのプログラミング言語で、数値処理能力が高く、機械学習の機能が豊富にある。1 年生の後学期の「AI 基礎」の授業で、学生全員が MATLAB をインストールしているため、2 年生の前学期の「データサイエンス基礎Ⅱ」でも MATLAB を活用することでスムーズに授業運営することができる考えた。また、本学では、入学したときに、学生全員が通常の性能のノートパソコンを購入しているため、学生のノートパソコンを用いて、通常の教室で時間内に実施可能な演習を開発した。機械学習の演習は、MATLAB の「Statistics and Machine Learning Toolbox」を活用した。データセットは、MATLAB にあるものばかりでなく、気象庁からダウンロードしたものを活用した。学生が自分の専門分野で機械学習を実施する際には、データセットをはじめに準備する必要がある。この授業では、気象庁からのデータセットをダウンロードし、学生自らデータを加工し、データセットを準備し、機械学習の教師あり学習による分類と回帰を行う演習を実践した。本学の特徴の一つの「プロジェクトデザインⅡ」、もしくは、卒業研究（「プロジェクトデザインⅢ」）で、人間には見つけにくい異なる視点の知見を見出すために、機械学習を活用する能力を養うことが非常に重要だと考えた。

表 1 で、第 1 回～第 7 回の各授業で実施する授業演習の主な内容を示した。これらの演習の他にも関連する内容の演習や解説も行っている。本学では、e シラバスを活用して対面授業、ハイブリッド授業、オンデマンド授業を実施している。ここで、e シラバスは、学生と教員がログインして閲覧できる学内の Web システムで、シラバス、授業目標、各授業回で実施する詳細な内容を閲覧することができる。授業に欠席した学生や、授業の予習・復習に活用で

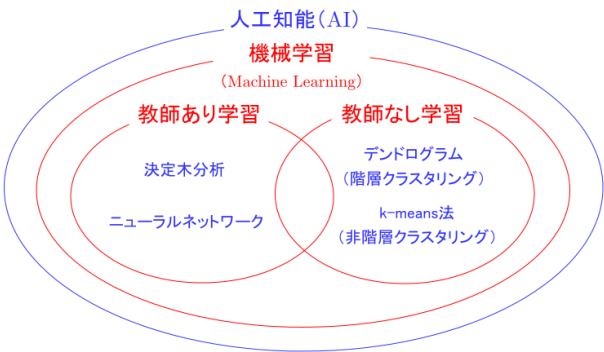


図 1 機械学習 (Machine Learning)

表 1 各授業で実施する授業演習

授業回	授業演習の主な内容	機械学習の種類
第 1 回	データセットの読み込みと機械学習による予測と検証	教師あり学習
第 2 回	階層クラスタリング：デンドログラム	教師なし学習
第 3 回	非階層クラスタリング：k-means 法	教師なし学習
第 4 回	決定木分析による分類	教師あり学習
第 5 回	決定木分析による回帰	教師あり学習
第 6 回	データセットの準備とニューラルネットワークによる分類と回帰	教師あり学習
第 7 回	最適なモデルの探索 (分類と回帰)	教師あり学習

きるように、すべての授業に対応した授業動画をクラウドの Box に準備し、e シラバスから閲覧できるようにした。学生は e シラバスから宿題とレポートのフォーマットとコードをダウンロードし、実施後に、PDF 形式で e シラバス上に提出する。教員は、その PDF を採点后、e シラバス上に返却した。学生が、採点された宿題とレポートをいつでも復習できるようにした。また、約 1,600 名の必修科目であるため e シラバスをすべてのクラスで共通化した。

2. 機械学習の教師なし学習

2. 1 階層クラスタリング：デンドログラム

第 2 回目の授業では、教師なし学習の階層クラスタリングの演習を実施した。データセットは、フィッシャーのアヤメの花の分析データ²⁾、座る・立つ・歩く・走る・踊るの 5 種類の動作に対する行動認識データ³⁾を活用した。

図 2 にアヤメの花のスケッチを示した。アヤメの花の花びらは 3 つで、がく片も 3 つある。表 2 は、フィッシャーのアヤメの花の分析データ「fisheriris.csv」²⁾の一部で、セトサ (Setosa)、バージカラー (Versicolor)、バージニカ (Virginica) の 3 つのアヤメの花のデータセットになる。データ数は合計 150 データあり、それぞれの種類ごとに 50 データずつある。機械学習の特徴量は、がく片の長さ、がく片の幅、花びらの長さ、花びらの幅 (表 2 の 1~4 列目)、正解データは、花の種類 (表 2 の 5 列目) になる。ここで、機械学習の教師なし学習は、特徴量だけから分類・回帰するモデル、機械学習の教師あり学習は、特徴量と正解データから分類・回帰するモデルを示す。

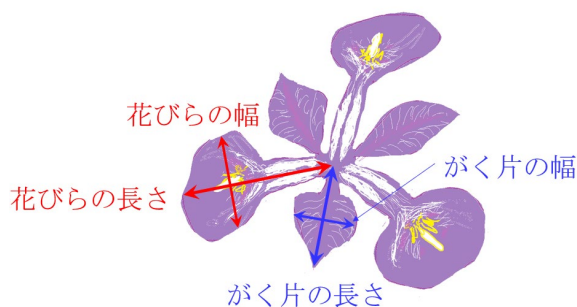


図 2 アヤメの花

表 2 アヤメの花の測定データ²⁾ (一部)

がく片の長さ	がく片の幅	花びらの長さ	花びらの幅	花の種類
5.1 cm	3.5 cm	1.4 cm	0.2 cm	セトサ
7.0 cm	3.2 cm	4.7 cm	1.4 cm	バージカラー
6.3 cm	3.3 cm	6.0 cm	2.5 cm	バージニカ

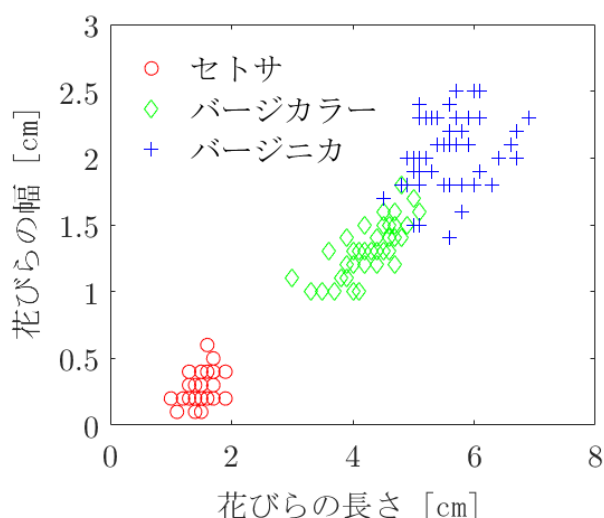


図 3 アヤメの花の散布図

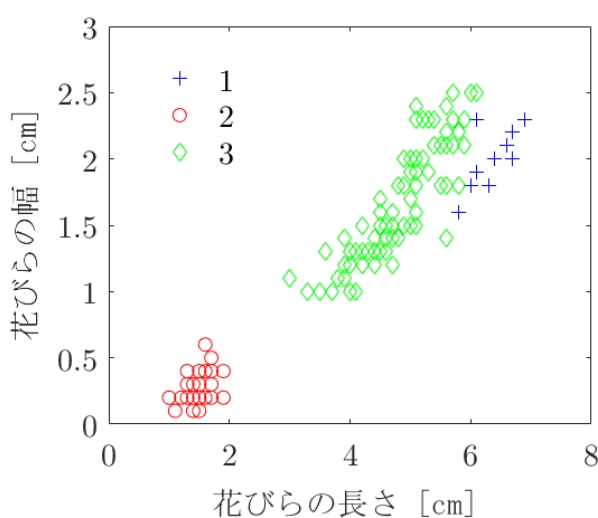


図 4 デンドログラムによる分類

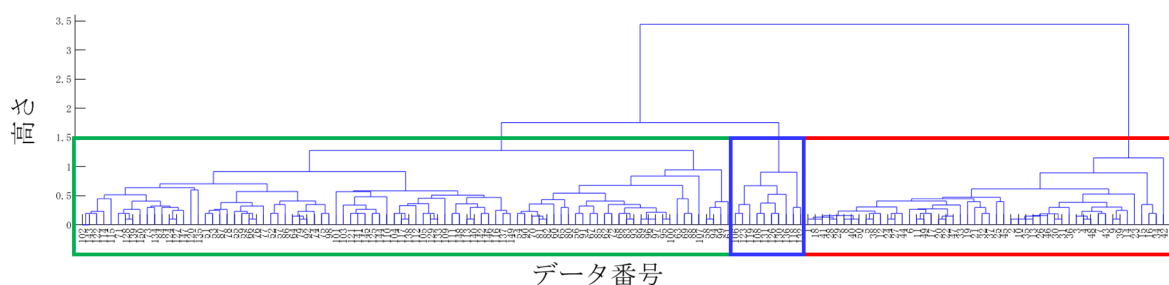


図5 デンドログラム

図3は、アヤメの花の観測データ（正解）の散布図を表す。横軸を花びらの長さ、縦軸を花びらの幅として、セトサを「○」、バージカラーを「◇」、バージニカを「+」で示した。図4は機械学習の教師なし学習の階層クラスタリングにより分類された散布図を示す。また、図5は、デンドログラムを示す。図5のデンドログラムにおいて高さ1.5で分類すると、緑枠、青枠、赤枠の3つの葉ノードに分類される。赤枠の葉ノードに含まれるデータ番号は1～50になり、図4の散布図の「○」で示した。ここで、データ番号1～50がセトサ「○」、データ番号51～100がバージカラー「◇」、データ番号101～150がバージニカ「+」に対応する。MATLABのコマンドウィンドウに、図5の3つの葉ノード（緑枠、青枠、赤枠）の中に含まれるデータ番号を表示して確認した結果、赤枠の中のデータ番号の個数は50個で、1～50のすべてのデータ番号が一つずつ含まれていた。したがって、セトサ「○」はデンドログラムにより適切に分類されたと考えられる。一方で、図5のデンドログラムの緑枠の葉ノードはデータ番号51～100ばかりでなく、データ番号101～150のデータもかなり含み、青枠の葉ノードは、データ番号101～150の一部のデータしか含まない。この結果は、セトサは100%適切に分類されたが、バージカラーとバージニカは適切に分類されなかったことを表す。分類の精度がよくないのは、教師なし学習の限界だと考えられる。

ここで、gscatter, clusterを用いたサンプルプログラムを参考にしながら、学生が自ら図4と図5のコード（mlxファイル）を作成する演習を行った。図4と図5において、通常のパフォーマンスをもつノートパソコンを活用して、150データを学習するのに要した時間は、それぞれ、1秒程度であった。さらに、5種類の動作に対する行動認識データ³⁾を活用したデンドログラムの演習も同様に行った。これらの結果を評価して、文章で考察させた。

2. 2 非階層クラスタリング：k-means 法

第3回目の授業で、教師なし学習の非階層クラスタリングの演習を実施した。ここで、k-means 法の特徴として、階層クラスタリングに比べデータ数が多くても分析することが可能で、計算速度がはやいことが挙げられる。5種類の動作に対する行動認識データ³⁾は、体に加速度センサをつけて、座る、立つ、歩く、走る、ダンスの5種類の運動をしたときの加速度の数値データになる。表3に、5種類の動作に対する行動認識データ³⁾の一部を示した。ただし、 x, y, z 方向の加速度の平均値を、それぞれ、 $\bar{a}_x, \bar{a}_y, \bar{a}_z$ [m/s²]とした。ここで、 x, y 方向は水平方向、 z 方向は鉛直方向を示す。

ここでは、座る、立つ、歩くの3つの動作の3つの加速度に着目し、合計2,491行のデータセットを準備した。

表3 5種類の動作に対する行動認識データ³⁾（一部）

動作	$\bar{a}_x$ [m/s ²]	$\bar{a}_y$ [m/s ²]	$\bar{a}_z$ [m/s ²]
座る	-0.314	0.684	-0.644
立つ	-0.087	1.005	-0.008
歩く	-0.189	0.960	-0.085
走る	-0.090	0.831	-0.131
ダンス	-0.042	1.008	-0.301

図6は、k-means法で分類するグループを3つにしたときの散布図マトリックスを表す。ここで、「×」はk-means法の3つのセントロイドを表す。図6の右下のヒストグラムで、青色のデータは座る動作に対応している。赤色のデータは歩く動作、緑色のデータは立つ動作に対応している。図6の左下の非対角成分は、3つのセントロイドによる分類を示す。ここで、加速度のデータの原点を正確に 0 m/s^2 に合わせて測定することは難しく、一般的にゼロ点補正が必要になる。歩く動作の z 方向の加速度の平均値 $\bar{a}_z [\text{m/s}^2]$ を 0 m/s^2 にするゼロ点補正により、立つ動作の $\bar{a}_z [\text{m/s}^2]$ は正の値になり自然に解釈することができる。さらに、ギャップ統計量、シルエット法など4つの手法を用いて、最適クラスター数を決定する演習も行った。図7は、シルエット法で、2乗ユークリッド距離を用いてクラスター分析した主成分分析の結果を示す。座る、立つ、歩くの3つの動作を分類した結果、最適クラスター数は2で、適切に3つに分類されなかった。

ここで、フィッシャーのアヤメの花の分析データ²⁾のk-means法による演習も実施した。学生は、4つの手法で最適クラスター数を決定する演習を通して、非階層クラスタリングについての理解を深めた。

3. 機械学習の教師あり学習

3. 1 決定木分析による分類と回帰

第4, 5回目の授業で、機械学習の教師あり学習として決定木分析の演習を行った。決定木分析は、人間の思考プロセスとよく一致し、分岐を可視化しているため、理解しやすい特徴をもつ。乱雑さ（不純度）を小さくして、情報利得を最大にするような学習をする。しかしながら、精度を上げるために、決定木の分岐を増やして過学習する傾向があり、そのため、分岐した枝を伐採して適度に分岐させる剪定を行う必要がある。図8で決定木分析による分類結果を示した。正解を「●」、不正解を「×」で示した。全データ150データの内、正解は145データになり、精度は96.7%であった。決定木（図9）で、はじめの質問の「花びらの長さが2.45 cm以上」で「No」を選択したときセトサになり、「Yes」を選択したときはバージカラーとバージニカのどちらかになる。

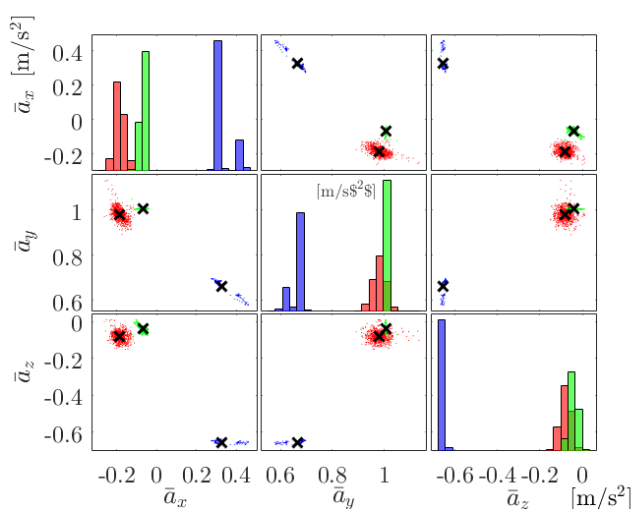


図6 k-means法の散布図マトリックス

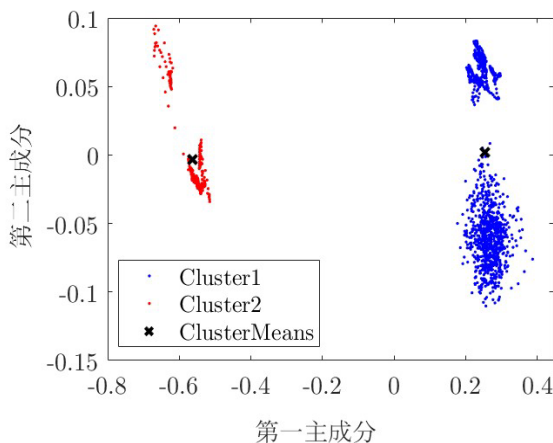


図7 k-means法（シルエット法）

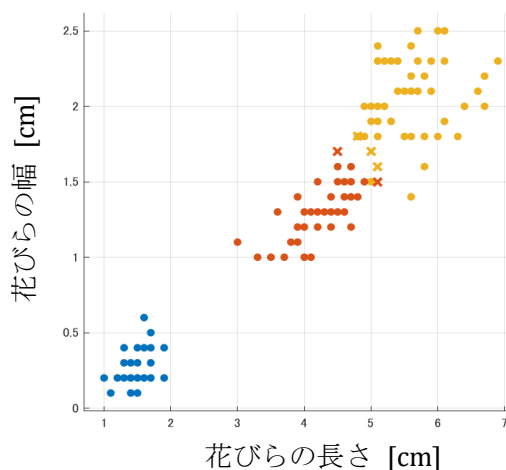


図8 決定木分析による分類

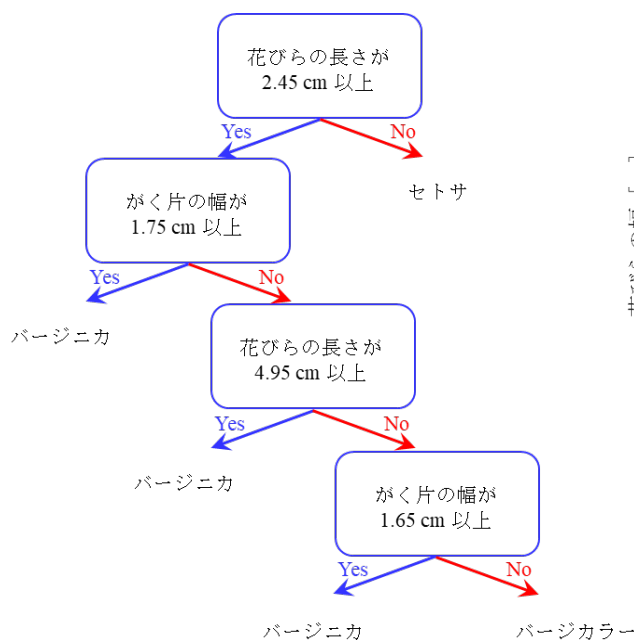


図 9 決定木

図 10 の 2.45 cm 未満（領域 1）の 50 データすべてがセトサ「○」として適切に分類された。続く質問の「がく片の幅が 1.75 cm 以上」、「花びらの長さが 4.95 cm 以上」で「Yes」を選んだとき、図 10 の右上の点線に囲われた枠（領域 2）で分類されたものが、バージニカ「+」になる。このように分類された結果、図 8 のように、145 データが正解になった。ここでは、k 分割交差検証（k-fold cross-validation）の分割数を 5, 10 と変えたときの精度（検証）を比較し、汎化性能の違いについて議論する演習も実施した。さらに、分類木ばかりでなく回帰木の演習も行った。図 11 は回帰木による箱ひげ図を示す。青い箱ひげ図は観測データ、橙色の箱ひげ図は決定木の回帰（予測）を表す。これらを比較するとよく回帰されている。学生は、決定木分析の分類と回帰の演習を通して、考察して理解を深めた。さらに、5 種類の動作に対する行動認識データ³⁾の決定木分析の演習も同様に実施した。

3. 2 データセットの準備

学生が自分の専門分野で機械学習を実施する際、自分でデータセットを準備する必要がある。どのようなデータを取得すれば、機械学習で分析することが可能になるか、どのようにデータセットを準備すれば機械学習ができるかについての手法を身につけることは非常に重要である。そのために、第 6 回目の授業では、学生が自ら気象庁⁴⁾から気象データ（CSV 形式）をダウンロードしたデータを加工し、データセットを自分で準備させた。表 4 は、加工後のデータセットを表す。1～3 列目の気温、降水量、降雪量の気象データを特徴量とし、4 列目の金沢市、札幌市、東京都、那覇市、京都市の City を正解データとした。その後、そのデータセットを活用し、ニューラルネットワークによる分類と回帰の演習を実施した。

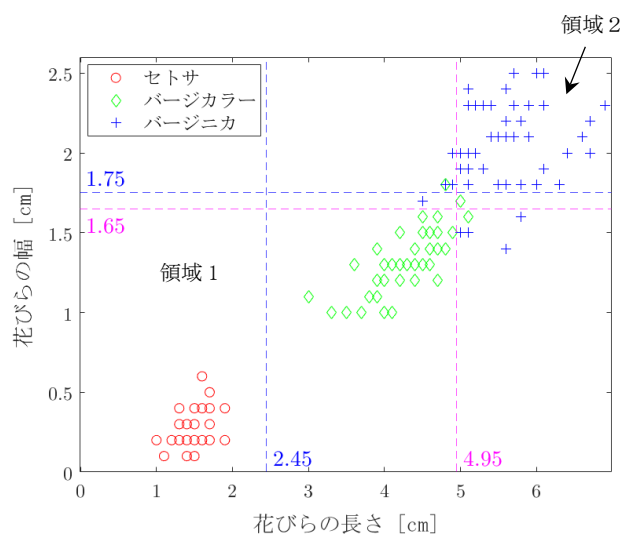


図 10 決定木による分類

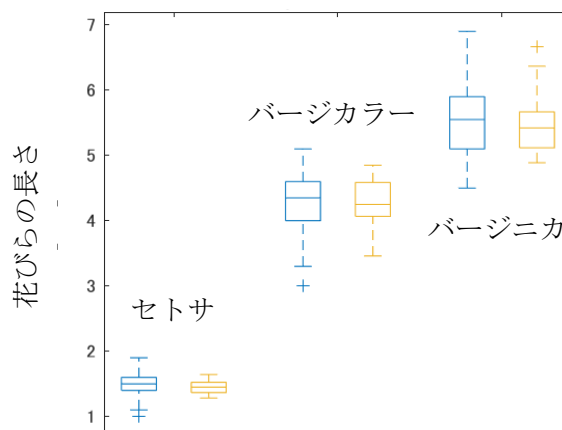


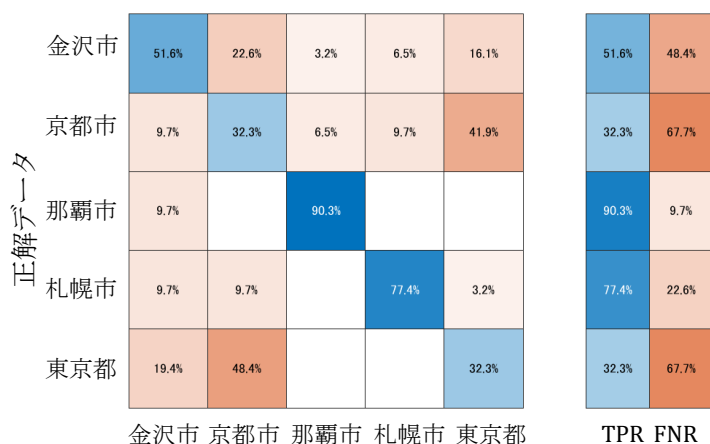
図 11 決定木による回帰

3. 3 ニューラルネットワークによる分類と回帰

ここでは、気象庁ホームページからダウンロードした気象データを用いて作成したデータセット（表4）を読み込み、機械学習の教師あり学習であるニューラルネットワークを活用して、アーキテクチャを変えた5つのモデルで機械学習を行った。図12は3層ニューラルネットワークの結果を示す。検証の混同行列において、横軸はニューラルネットワークによる予測、縦軸は正解データ（観測データ）を表す。対角成分はニューラルネットワークによる予測が正解した割合を示し、非対角成分はニューラルネットワークによる予測が不正解になった割合を示す。

表4 気象庁ホームページ⁴⁾からダウンロードした気象データを用いて作成したデータセット（一部）

Temperature [°C]	Rain [mm]	Snow [cm]	City
6.5	18.5	0	Kanazawa
3.6	4.5	0	Kanazawa
2.9	3	0	Kanazawa
2.2	7	0	Kanazawa
2.9	6	1	Kanazawa



ニューラルネットワークによる予測

図12 検証の混同行列

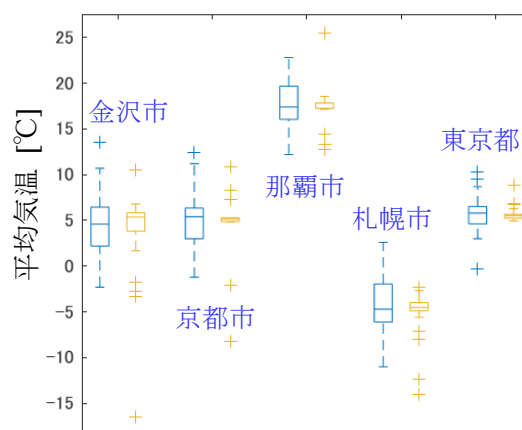


図13 ニューラルネットワークによる回帰

混同行列の右側の図は、真陽性率：TPR（True Positive Rate）と偽陰性率：FNR（False Negative Rate）のそれぞれの値を表す。3つの気象条件（気温、降水量、降雪量）のみから、都市を分類する学習において、那覇市が90.3%の精度で正解し、札幌市が77.4%の精度で正解した。金沢市と京都市と東京都は、気象条件が似ていたため精度が低かった。

図13は、ナローニューラルネットワークによる回帰の結果を示す。青い箱ひげ図は観測データ、オレンジ色の箱ひげ図は回帰を示す。ナローニューラルネットワークの二乗平均平方根誤差（RMSE）3.3351となり、中央値はよく予測されているが、第1四分位数、第3四分位数に関しては、予測がよくされていない。今後、気象データの特徴量を工夫する必要があると考えられる。

3. 4 機械学習の最適なモデルの探索

第7回授業では、機械学習の教師あり学習の最適なモデルを探索する演習を行った。はじめに分類学習の最適なモデルについて探索した。5種類の動作に対する行動認識データ³⁾の3,355データを抽出し、データセットを準備した。図14は、 x 軸、 y 軸、 z 軸を、それぞれ、 $\bar{a}_x, \bar{a}_y, \bar{a}_z$ [m/s²]として、

座る「▽」、立つ「△」、歩く「□」、走る「・」、ダンス「☆」のデータになる。ここで、 z 軸方向は鉛直方向を示す。図 14 において、走る「・」とダンス「☆」のデータがかなり重なっていたため、教師あり学習の精度を検証するのに適していると考えた。実際の演習では、31 個の機械学習の教師あり学習のモデルで分類学習を行い（計算時間 1 分 30 秒程度）、最適なモデルを探索し、学生同士で議論させた。31 個のモデルの内、最も高い精度は 95.3 %で、細かいガウスのサポートベクターマシン：Support Vector Machine (SVM) のモデル（計算時間 3.52 秒）になった。図 15 の検証の混同行列から、動作の精度は、座る動作は 100 %、立つ動作は 99.8 %、歩く動作は 97.0 %となり、これらは高い精度で分類された。一方で、走る動作の精度は 72.8 %で、18.0 %はダンス、8.9 %は歩く動作に間違えて分類された。

つぎに、回帰学習の最適なモデルについて探索した。機械学習の教師あり学習の 26 個のモデルで、フィッシャーのアヤメの花の分析データ²⁾の花びらの長さを回帰する学習を行い、最適なモデルを探索し、学生同士で議論させた（計算時間 40 秒程度）。26 個のモデルの内、二乗平均平方根誤差 (RMSE) が最も小さい値は 0.236 で、ステップワイズ線形回帰モデルであった（計算時間 6.86 秒）。ここで、分類学習も回帰学習の最適なモデルは、学生間で異なることもあるため、最適なモデルについて学生間で議論し、知見を深めた。

4. 学生の達成度

図 16 と図 17 は、1 クラスの 24 名（アンケート同意者）の結果になる⁵⁾。図 16 で、データサイエンスの演習がとても楽しかった、楽しかったと回答した学生の割合は 91.7%，機械学習を実施する能力がとても身に付いた、身に付いたと回答した学生の割合は 87.5%で、否定的な回答をした学生は一人もいなかった。以下のようなコメントがあった。

- MATLAB は意外と扱いやすかった。今後の授業や活動に役立てていきたい。
- PD や専門実験などのデータ分析などに機械学習を活用していきたい。
- MATLAB を使えるようになった。PD などでも使うようにしたい。

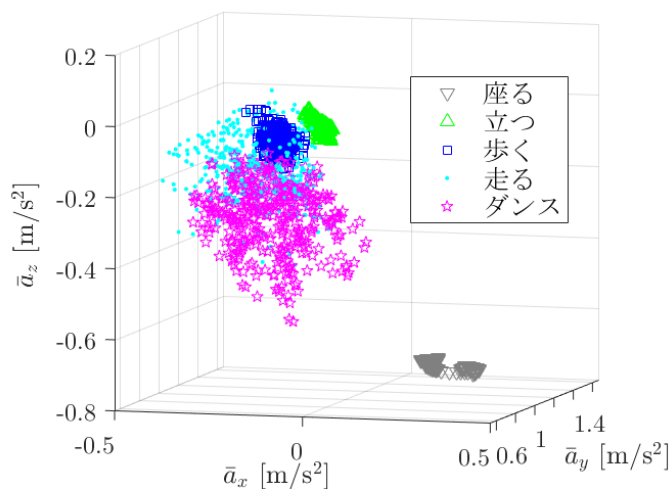


図 14 5つの動作のデータ

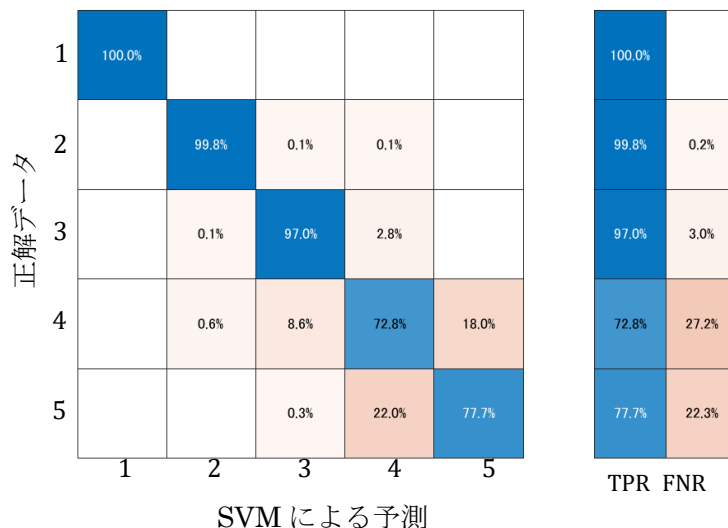


図 15 検証の混同行列

○ 今後の「PDⅢ」や夢考房の課外活動などでセンサを扱うときに MATLAB を使用したいと考えているので、またご連絡させていただき、使用方法を教えていただきたい。

データサイエンスの演習を楽しく感じることで、授業後にも機械学習を自分で実施するきっかけをつくるのが授業の狙いであったが、その目的はある程度達成できたと考える。

図 17 は、達成度に関する設問になる。「機械学習の教師あり学習の最適なモデルを検証し、分類と回帰を実施することができる。」に関して、達成度 80%以上の学生の割合は 82.6%になり、達成度 60%未満の学生は一人もいなかった。「データセットを自分で作成した上で機械学習を実施することができる。」に関して、達成度 80%以上の学生の割合は 66.7%になり、達成度 40%未満の学生はいなかった。以下のようなコメントがあった。

○ 回帰分析、決定木分析を MATLAB で扱う方法を学んだ。またデータから考察する能力を養うことができたと思う。

○ 様々なデータを分析して最適なモデルを見つけ出すことができるようになった。データセットを用意することや様々な能力が身についたと思う。

各自の専門分野のデータから、学生が自らデータセットを準備し、最適なモデルで分析する能力を身につけることも授業の狙いの一つとしていたが、図 17 のアンケート結果と自由記述から、この目的もある程度達成できたと考える。

また、授業では、AI や機械学習の最新的话题について取り上げた。この点に関して、以下のようなコメントもあった。

○ MATLAB の正しい使い方を学ぶと同時に、データの取り扱いや重要情報についての細かいところも知ることができた。

その一方で、MATLAB を用いた機械学習について、以下のようなコメントもあった。

○ MATLAB を使った機械学習についてまだよく理解していないため、今後しっかりと理解し、使いこなせるようにしていきたい。

○ 完璧とは言えずとも、ある程度自分で分析して説明できるようになった。

機械学習の理解度が、十分でない学生もいたため、今後、受講者全員が機械学習の理解を深めることができるように工夫する必要性がある。

データサイエンスの演習は楽しかったですか。



■ とても楽しかった。 ■ 楽しかった。 ■ どちらでもない。
■ 楽しくなかった。 ■ 全く楽しくなかった。

機械学習を実施する能力は身に付きましたか。



■ とても身に付いた。 ■ 身に付いた。 ■ どちらでもない。
■ 身に付かなかった。 ■ 全く身に付かなかった。

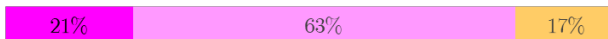
図 16 アンケート結果

■ 100% ■ 80% ■ 60% ■ 40% ■ 20% ■ 0%

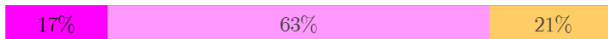
クラスター分析を活用して分析することができる。



決定木分析を活用して分析することができる。



ニューラルネットワークを活用して分析することができる。



データセットを自分で作成した上で機械学習を実施することができる。



機械学習の教師あり学習の最適なモデルを検証し、分類と回帰を実施することができる。



各自の専門分野におけるデータサイエンスの意義について基本的な説明ができる。



図 17 学生の達成度

5. まとめ

2023 年度に新規開講した「データサイエンス基礎Ⅱ」について、授業設計と学生の達成度について解説した。この授業では、データサイエンスの手法として機械学習を用いて演習をおこなった。機械学習の教師なし学習として、階層クラスタリング（デンドログラム）と非階層クラスタリング（k-means 法）の分類をおこなった。ここで、k-means 法では、4 つのモデルで最適クラスター数を検証した。また、教師あり学習として、決定木分析、ニューラルネットワークなどのモデルにより分類と回帰をおこなった。さらに、31 個のモデルで分類を行い、26 個のモデルで回帰を行うことで、最適なモデルについて探索して理解を深めた。

学生が各自の専門分野で、機械学習を実施するためには、データセットを自分で準備する必要がある。そのために、この授業では、気象庁からダウンロードしたデータを自分で加工し、特徴量と正解データを入れたデータセットを自分で準備する演習を行った。その後、自分で準備したデータセットにより、実際に、機械学習の教師あり学習で分類と回帰を行った。このような経験を通して、専門分野で機械学習を実施できる基本的な能力を育成できると考える。

学生のアンケート結果では、9 割程度の学生がデータサイエンスの演習が楽しかったと回答し、機械学習の能力が身に付いたと回答していたため、授業の狙いはある程度達成できたと考える。

今後も機械学習の理解度をさらに深めるように授業を改善する。

参考文献

- 1) 「AI 戦略 2019」の概要と取組状況, 令和元年 11 月, 内閣府 政策統括官 (科学技術・イノベーション担当), URL: <https://www5.cao.go.jp/keizai-shimon/kaigi/special/reform/wg7/20191101/shiryoul.pdf> (参照日: 2024 年 8 月 19 日)
- 2) MathWorks 社の「Statistics and Machine Learning Toolbox」のサンプルデータセット「fisheriris.csv (fisheriris.mat)」, URL: <https://jp.mathworks.com/help/stats/sample-data-sets.html> (参照日: 2024 年 8 月 19 日)
- 3) MathWorks 社の「Statistics and Machine Learning Toolbox」のサンプルデータセット「humanactivity.mat」, URL: <https://jp.mathworks.com/help/stats/sample-data-sets.html> (参照日: 2024 年 8 月 19 日)
- 4) 気象庁ホームページ, 各種データ・資料「過去の地点気象データ・ダウンロード」, URL: <https://www.data.jma.go.jp/gmd/risk/obsdl/index.php> (参照日: 2024 年 8 月 19 日)
- 5) 人を対象とする研究倫理審査の承認番号「2402008」, 研究倫理委員会, 金沢工業大学.

[受付日 令和 6 年 (2024 年) 8 月 19 日, 採択日 令和 6 年 (2024 年) 10 月 22 日]



工藤 知草
准教授・博士 (理学)
基礎教育部
数理・データサイエンス・
AI 教育課程
数学教育・物理教育・情報
教育・量子力学



谷口 進一
教授・博士 (学術)
基礎教育部
数理・データサイエンス・
AI 教育課程 主任
質的研究・効果測定・数理工
基礎教育