

数理・データサイエンス・AI教育の紹介

# 和歌山大学の数理・データサイエンス・AI教育プログラム 「データサイエンスへの誘い（応用基礎レベル）」の紹介 ～文理融合実践型DS教育における学生の受講動機向上のための工夫～

和歌山大学  
データ・インテリジェンス教育研究部門部門長  
システム工学部教授

和歌山大学  
データ・インテリジェンス教育研究部門講師

和歌山大学  
データ・インテリジェンス教育研究部門講師

吉野 孝

西村 竜一

三浦 浩一



(左から 吉野、西村、三浦)

## 1. はじめに

2023年の報告<sup>[1]</sup>では、本学での数理・データサイエンス・AI教育の取組みについて、その全体を示すとともに、認定を受けたばかりの文部科学省「数理・データサイエンス・AI教育プログラム認定制度」のリテラシーレベルプラスについて紹介いたしました。そのなかで、本学は、2023年度に応用基礎レベルに申請予定と述べていましたが、幸いにもその後、応用基礎レベルプラスの認定を受けることができました。本学は、全学を対象に、リテラシーレベルプラスと応用基礎レベルプラスの認定をともに受けることができましたこととなります。

そこで、今回は、応用基礎レベルに該当する科目を中心に筆者たちの取組みを紹介させていただきます。なにをどのようにすることで、全学を対象とした応用基礎レベルのプラス認定を受けることができたのか、数理・データサイエンス・AI教育プログラムにこれから取り組む、あるいは、改善を考えている担当者さまの一助になれば幸いです。

なお、本学は、教育学部、経済学部、システム工学部、観光学部、社会インフォマティクス学環(2023年度開設)の4学部1学環で構成されている地方国立大学です。工学系の学部があり、我々の教育プログラムは文理融合をキーワードの一つにしていますが、キャンパス内は、いわゆる文系色が強めだと思えます。特に同じような規模感の大学のみなさまに参考にしていただくことを想定しています。

## 2. 「データサイエンスへの誘い」プログラム

まず、筆者たちの教育プログラムの全体構成についておさらいをします。図1のカリキュラムツリーをご確認ください。文献[1]に掲載した図と同じものです。

本学は、2021年度にリテラシーレベル、翌2022年度には、リテラシーレベルプラスに認定されています。リテラシーレベルを構成しているのは、1年次1Q「データサイエンスへの誘いA」と2Q「データサイエンスへの誘いB」の2科目のみです。応用基礎レベルプラスの認定を受けたのは2023年度です。構成科目は、大学院科目を除いた「データサイエンスへの誘いA/B」を含むすべての科目です。「誘いA/B」は、リテラシーレベルと応用基礎レベルの両方に含まれます。

「誘い」は「いざない」と読みます。すべての学生にとって馴染みがあることから、導入の科目名から名前をとって、本学のプログラム全体を

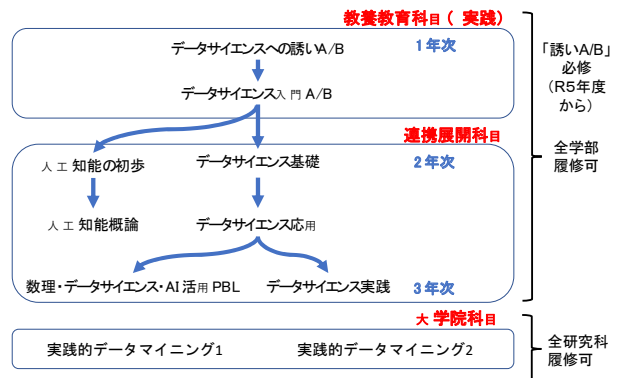


図1 本学の教育プログラムの体系

「データサイエンスへの誘い」プログラムと名付けました。

本学では、本プログラムを2019年度に立ち上げました。当初からの科目は、「データサイエンスへの誘いA/B」「データサイエンス入門A/B」「データサイエンス基礎」(2020年度～)「データサイエンス応用」(2020年度～)「データサイエンス実践」(2021年度～)「実践的データマイニング1」「実践的データマイニング2」です。「人工知能の初歩」「人工知能概論」「数理・データサイエンス・AI活用PBL」は、後から追加しています。筆者たちが幸運だったのは、この全学対象の科目群をゼロから設計できたことです。以前からの既存科目は組み込んでいません。

### (1) 履修者数・履修率の向上に向けた工夫

「誘いA/B」は、2023年度から全学必修科目です。毎年度1,000人程度の学生が履修します。ただし、それ以外は、卒業に必要な単位に認められる正規科目ですが、選択科目です(一部、社会インフォマティクス学環では必修科目)。このため、多くの学生にとっては、本プログラムは卒業に必要ではなく、履修生の数を確保することが難しくなっています。

また、「誘いA/B」と「入門A/B」は、教養教育科目ですが、他は「連携展開科目」という科目群の扱いになっています。これは、教養科目と学部専門科目の接続を目的とした全学対象の科目群です。ただし、「連携展開科目」の各学部での扱いは異なっており、基本的には「自由選択科目」として扱われるため、卒業要件における優先度は高くありません。

このため、本プログラムを維持するためには、履修者数を増やすための工夫が必要になります。もちろん、内容を工夫して、魅力を高めることで履修者を増やすことが本質ではあります。しかし、並行して、履修がしやすい制度や環境も必要となります。

本学では、コロナ禍前から、動画やビデオ会議サービスを用いた遠隔授業(オンライン授業)の推進及びBYOD PCの導入を必須化してきました。そのため、本プログラムでは、動画を用いたオンデマンド型遠隔授業を中心とした設計を当初から行っています。具体的には、「誘いA/B」「入門A/B」「基礎」「応用」は、動画を用いたオンデマンド型遠隔授業を基本としています。本学では、オンデマンド型授業であっても、時間割の中に開設時間を配置する必要があります。本プログラム

では、オンデマンド型の科目の開設を18:10～19:40の6時限に設定しています。このため、他の科目と時間割上の競合はありません。また、教室などの物理的制約を受けないことから、定員などの受講制限を排除しています。教員の負担は大きくなりますが、そこは運用で補うこととし、まずは、多くの学生を積極的に受け入れる姿勢を示すことが大切であると考えています。

また、「基礎」「応用」「実践」「数理・データサイエンス・AI活用PBL」は、システム工学部においては、学部の専門教育科目として単位が認定されるようになってきました。これも受講生の確保を狙った工夫の一つです。本プログラムの開始時に、学部教務委員会と協議して実現することができました。

### (2) 受講生サポートの充実

詳細は、前回の報告[1]に記載しましたので省略しますが、遠隔授業が多くなるため、受講生サポートの充実を欠かすことはできません。学生からの質問などに半自動で対応するために、独自のLINEアプリの開発を行っています。また、週に1回、他の授業と競合しない18:10以降にTeamsビデオ会議を用いたオンラインサポート室を開設しています。このオンラインサポート室は、本プログラムの科目で共通であり、すべての科目の受講生が参加できます。学年や科目を超えた学生同士が交流できる、言わば本プログラム全体のホームルームになることを期待しています。学生同士での問題解決を促し、結果として教員や大学院生TAの負担を減らしながら、サポートの充実につながることを期待しています。

### (3) デジタルオープンバッジと物理バッジ

デジタルオープンバッジの導入による単位修得状況や学びの可視化も工夫です<sup>1)</sup>。一般財団法人オープンバッジ・ネットワークのシステムを利用して、2022年からオープンバッジを発行しています。オープンバッジは、スマホのアプリに配信されて、オンラインでコレクションすることができます。

本プログラムでは、「科目バッジ」の他に「単位積み上げ型バッジ」を発行しています。「単位積み上げ型バッジ」は、所定の4単位を取得すると「ブロンズレベル」が与えられます。ブロンズレベルのあとに、「データサイエンス基礎」を修得すると、「シルバーレベル」、さらに、シルバーレベルのあとに、「データサイエンス応用」を修

得すると「ゴールドレベル」を獲得することができます。ゴールドレベルを獲得したあとに、「データサイエンス実践」か「数理・データサイエンス・AI活用PBL」の単位を修得できると、「ダイヤモンドレベル」になります。

本プログラムの応用基礎レベルの修了要件は、この「ダイヤモンドレベル」と一致させています。つまり、オープンバッジで、本プログラムでのサートیفिकーションが完結できています。学生は、達成レベルをスマホで常に確認できます。また、運用も効率化できることから、この枠組みは、リーズナブルであると感じています。学生からも好評で、モチベーション維持に一定の効果があるようです。学生の気持ちに伝えるべく、2023年度からオープンバッジと同じデザインの実体のある物理バッジ(写真1)の発行をはじめました。ダイヤモンドレベルのみですが、本物のバッジを達成者に授与しており、セレモニーとして授与式も開催しています(写真2)。



写真1 物理バッジ



写真2 「ダイヤモンドレベル」授与式

### 3. 「データサイエンス基礎」

制度の話が長くなってしまいましたが、ここからは、応用基礎レベルの中心的な科目である「データサイエンス基礎」「データサイエンス応用」「データサイエンス実践」の中身を紹介いたします。

本プログラムでは、「誘いA/B」では主にExcel、「入門A/B」ではRを使いますが、「基礎」からはプログラミング言語にPythonを使用しています。

「データサイエンス基礎」(2年次前期 Semester)は、このあとの科目の実践的学習に必要な導入と基本知識・技術の獲得フェーズに該当します。Pythonプログラミング、統計学や数学の基本的知識、様々な分析手法の理解に加えて、それらの利用方法を身に付けることができます。認定制度の応用基礎コア「I. データ表現とアルゴリズム」は、「データサイエンス基礎」のみでも最小の条件を満たすように設計しているため、「基礎」は、応用基礎レベルの必須科目です(他の科目を組み合わせることでさらに多くの範囲をカバーす

ることができる)。

授業は、Pythonの実行環境として、自分のBYOD PCにJupyterをインストールするところから始まります。本学では全学共通のJupyterサーバを運用しており、受講生は自由に利用できます。しかし、データサイエンスの学習に取り組む上で、自分で実行環境を準備する行程は必要だと考えます。また、このときにファイル操作などの基本的なスキルの不足を確認することができます。本学では、BYOD PCのOSは、Windowsに統一されていますが、それでも個々の環境は異なるため(あえてMacを購入して、使いこなせない学生も多い)、環境構築の際には様々なトラブルが発生します。オンラインや対面でのサポートを行いながら丁寧に対応しています。

環境構築を行った後は、Pythonを使ったデータの操作や加工、分析のための可視化手法を学びます。Pandas、NumPy、Matplotlibといった必須ライブラリを中心に基本的な使い方を学びます。

次に、確率論、推測統計、仮説検定の基礎知識を学びます。統計学の理解は、データ分析の結果を解釈するために必要不可欠です。

後半では、機械学習アルゴリズムの理解と活用方法を取り扱います。回帰分析、分類、クラスタリング、次元削減など、基本的なアルゴリズムを学び、PythonのScikit-learnライブラリとサンプルデータを用いて使用方法を学習します。また、機械学習モデルの評価方法についても学びます。

「データサイエンス基礎」では、難易度を段階的に上げることで、受講生が無理なく理解を深めることができるように注意しています。次の授業である「データサイエンス応用」にスムーズにつながるように、基礎の確立を図っています。また、プログラミングの際は、エラーなど様々な不具合が起きるのが常ですが、対処方法だけでなく、原因の深掘りを促すような指導を心掛けています。問題解決能力の基礎をしっかりと身につけることで、その後の実践的な学習への移行が可能となっています。

### 4. データサイエンス応用

「データサイエンス応用」(2年次後期 Semester)では、テキストマイニングをテーマとしています。テキストからの特徴抽出やテキスト分類等の技術を理解し、実際に利用することで、SNSや新聞記事、書籍、論文等のデータから発見を得るための手法を身に付けることを目指しています。関連して、音声言語やビッグデータ、Web API等

を活用するための技術について、体験的に学びます。

「データサイエンス応用」は、認定制度の応用基礎コア「Ⅱ. AI・データサイエンス基礎」の必須科目となっており、この授業だけで同コアの最小要件を満たすことができます。また、後述する小論文課題などは、応用基礎コア「Ⅲ. AI・データサイエンス実践」の範囲も含んでいます。

Pythonでプログラミングをすることになりますが、受講生がゼロからコーディングすることは困難であるため、担当教員が用意したサンプルコードを実行し、動作確認したあとに、与えられた課題に沿って、そのコードを改造することで学習を進めることにしています。その際には、参考書やウェブ資料などで調査をし、自己解決ができるようになることを求めています。翌週には、課題を解決した動作するコードの例を動画解説付きで公開していますが、これは単純に正解を提示するだけのものではありません。自分が作成したコードと見比べることで、より良い解決方法を考えるきっかけとしてもらうことを強く意識しています。

前半の内容としては、テキスト処理のための正規表現、形態素解析、クレンジング、特徴量、ベクトル表現、言語モデルなどになります。データには、ウェブサイト「青空文庫」から取得した著作権切れの小説を使用しています。例えば、作家や年代、ジャンルによって異なる小説中の傾向などを分析します。機械学習や簡単な深層学習を導入して、テキスト分類の実験も行います。

後半では、世の中に存在する各種データを探索し、それらを利用したデータ利活用の実験を実施します。そのために、分析対象とするデータの収集と整備からはじめます。ウェブクローリングやスクレイピング、構造化データ・半構造化データの取り扱いなどを学びます。また、音声データからテキストを取得する手段として、音声認識技術を使用することもあります。

データの取得の際には、著作権等の理解が必要です。データ分析に日本の著作権は寛容ですが、著作権者やウェブサイト管理者等に配慮して、ルールやマナーを守った実験が求められます。そのために必要な技術についても取り扱います。

ビッグデータのエンジニアリングの体験を意図して、Wikipediaダンプデータも使用しています。しかし、学生のBYOD PCではリソースが足りません。全学Jupyterサーバを利用してトラブルの事前回避を図っていますが、提出直前にはサーバがダウンしてしまいます。安定的な授業のため

の環境改善が、今後の課題の一つだと言えます。

この授業のもう一つの特色に最終課題としている小論文があります。収集データを分析し、4ページ程度の小論文にまとめる課題です。データサイエンスの業務では、図表等を用いて結果をレポートにまとめる行程が大切になります。学生には卒業論文の練習にもなります。結果としては、以下のような驚くほど良質なレポートがありました。

- レビューを用いた大学生協アプリ(公式)の問題点の分析
- SNSの誹謗中傷を表す文の分析及び予測
- 商品検索APIを用いた食品商品名の分析
- 検索数および検索トレンドから読み取る地域への影響力と人々の関心の度合いについて
- 歌詞の特徴抽出による作詞者ごとの楽曲分類

小論文の質をよくするために、ここでも一つの工夫をしています。小論文の下書きの提出を最終回の1回前の課題としています。そして、最終回の授業では、受講生の前で、教員が一人ずつ下書きを読んで口頭でコメントをしています。コメントは、内容や書き方の助言です。分析が不足している場合には、考え方のヒントを与えたりもします。本人が参考にするだけでなく、他人向けのコメントも知ることができます。最終課題は、コメントを反映させた小論文(清書)の提出です。

「データサイエンス応用」は、基本はオンデマンド型遠隔授業ですが、この最終回(コメントライブと呼んでいます)は、Teams会議をつかった同時双方向型で実施しています。見やすい手元の画面で他の受講生の小論文を参考にできる。それに対しての教員のコメントも聞くこともできる。コメントライブは今後も続けたいと考えています。

## 5. 「データサイエンス実践」

本プログラムでは、応用基礎コア「Ⅲ. AI・データサイエンス実践」の内容を含む授業科目として、前述の「データサイエンス応用」は必須、そして、「データサイエンス実践」と「数理・データサイエンス・AI活用PBL」のどちらか一つを選択必須としています。「数理・データサイエンス・AI活用PBL」は、他大学と共同開催している科目であり、毎年内容が変化するため、本稿では、「データサイエンス実践」についてのみ紹介をさせていただきます。

「データサイエンス実践」(3年次1Q)は、その名の通り、応用基礎コア「Ⅲ. AI・データサイエ

ンス実践」に対応する科目です。人や社会にかかわる具体的な課題の解決に、データを活用できる能力を育成することをねらいとしています。本プログラムを応用基礎レベルに申請する際、最初からプラス認定を目指した根拠となった授業でもあります。

この授業では、地元大手企業から提供を受けた本物のPOSデータ(レジでの支払時等に記録される商品の購買データ)を用いて、3～4名のグループワークを実施、企業が実際に抱える課題の解決を提案します。実際の店舗を調査するなどし、机上の理論のみではない、実践的な体験をすることになります。

途中には、経過報告の報告会を頻繁に実施しており、データから導き出される提案の着実な具体化作業をサポートしています。さらに、「課題の発見と定式化」「データの取り扱い」「モデル化」「結果の可視化」「検証、活用」等のデータサイエンスの活用に必要な一連のプロセスの理解を深めます。

整備が都合よく完了していないリアルなデータを使用しており、必要な前処理等の労力を実感することができるため、実際のプロセスに必要な「手触り感」の体験も含めた学修を行うことができます。

企業の担当者が、ほぼ毎回の授業に参加しており、現場目線のコメントがもらえることも特色です。写真3に最終発表会の発表の様子を示します。

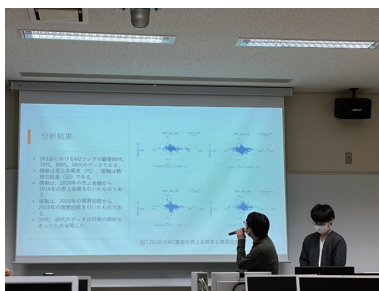


写真3 最終発表会（データサイエンス実践）

これまで、地元大手企業から提案された具体的なテーマとしては、次のようなものがありました。

- 環境変化による購買変化
  - 温暖化等災害の増加
  - コロナ禍から現在までの購買変化
  - 値上げ
- 次世代食生活への対応
  - 次世代ミートなど
- 会員施策への提案
  - 会員数を増やす工夫

教員の目線になりますが、流通業の課題は、学生に身近であり、また、様々な観点から社会の観察が必要になるため、データサイエンスを学ぶ題材としては、かなり良質なものだと考えています。

## 6. これからの取組み

本稿の執筆時点ではまだ途中ですが、我々の新しい取組みを2つ紹介いたします。

- データサイエンス教育では、教材として使用するデータが重要です。我々は、POSデータを使用するなどしてリアルなデータにこだわっています。学生に興味を持ってもらい、学生の履修動機を高めるために、学生自身に関するデータを集めて教育用に整備する予定です。これまでにポイント配布型アンケートを用いて、大学生の日常や習慣、関心がある事柄等を調査しています。
- 2024年12月、和歌山県・白浜で2泊3日のデータサイエンスハッカソンを地域企業等と協力して実施する予定です。他大学の学生グループも参加いただけます。単位が認定される授業ではありませんが、世界で活躍することができるハイレベル人材を育成するためのイベントとして成長させたいと考えています。

## 7. おわりに

本稿では、本学の数理・データサイエンス・AI教育の中で、特に応用基礎レベルに関連する取組みを紹介させていただきました。

我々は、今後も、より魅力的な教育を展開したいと考えています。今後ともご助言等いただけますようお願いいたします。特に、本学は、文部科学省「数理・データサイエンス・AI教育の全国展開の推進」において、社会科学の特定分枝校に選定されています。本稿をきっかけとし、社会科学分野等の皆さまとご意見の交換等をいただけますと幸いです。

## 参考文献および関連URL

- [1] 吉野 孝, 西村 竜一, 三浦 浩一, “和歌山大学の数理・データサイエンス・AI教育プログラム～実践的教育を軸とした文理隔たりのない体系的な取組み～”, 大学教育と情報 (公益社団法人私立大学情報教育協会), 2022年度 No.4, 通巻181号  
[https://www.juce.jp/LINK/journal/2303/pdf/03\\_02.pdf](https://www.juce.jp/LINK/journal/2303/pdf/03_02.pdf)