

数理・データサイエンス・AI教育の紹介

早稲田大学における全学データ科学教育プログラム

早稲田大学
データ科学センター所長・理工学術院教授

松嶋 敏泰



早稲田大学
データ科学センター教授

野村 亮



(左から 松嶋、野村)

1. 概要

本学では2018年度より全学向けのデータ科学教育プログラムを提供しています。20科目を越える科目からなり、その提供にあたり履修目標を明確にするためにリテラシー級から上級まで本学独自に4つの級を用意しています。その中のリテラシー級および初級の科目群が文部科学省の「数理・データサイエンス・AI教育プログラム認定制度 (MDASH)」の「リテラシーレベル」および「応用基礎レベルプラス」の認定を受けました。本稿では、本学の提供するデータ科学教育プログラムの概要とその特色について述べていきます。

2. 本学データ科学教育プログラム

本学は、「高度なデータ分析能力」を持った研究者と様々な領域における「深い専門知識」を持った研究者の融合によるデータ駆動型の最先端の研究を推進するとともに、専門知識にデータ科学を活用して新しい知見を創出できる実践的な人材の育成を目的として、2017年12月に「データ科学センター (以下DSセンター)」を設置しました。

図1はDSセンターの学内向け機能を示しています (学術院とは学部と大学院を合わせた組織です)。縦方向は各学術院で学ぶ専門性を示しており、それらに対して横車を刺すように全学に向けたサービスを展開しています。図から分かるようにDSセンターは本学全学に対するデータ科学教育・研究の両面を推進する役割を担っており、主に低学年次生に対しては教育プログラムを提供し、高学年あるいは大学院生に対しては卒業論文や学会発表などの研究レベルでのデータ科学活用に関する相談を受け付けています。これらはいず



図1 本学データ科学センターの概要

れも本学の学生であれば学部や研究科を問わず誰でも受けることのできるサービスという位置づけです。このような役割を担っているDSセンターの目標は、本学に所属する全学生のデータ科学力向上、さらにこれに伴う自身の専門領域における研究力強化にあります。以上を視野に入れてDSセンターでは学内向けに下記のような教育プログラムを複合的に展開しています。

- 正規授業科目
- 各種セミナー・ワークショップ
- データサイエンスコンペティション
- インターンシップ
- キャリア教育プログラム
- 大学院生用自学自習コンテンツ

これらは内容は独立していますが、関連を持っています。例えば、正規授業科目でデータ科学に関して一定の知識を学んだ後に、データサイエンスコンペティションや各種インターンシップなどでより実践的にデータ科学の活用方法について学んだり、データ科学の入門的なセミナーから正規授業科目への誘導を図るなどで、各種コンテンツの連携により教育プログラムとして効果的に全学的なデータ科学力向上を目指しています。

以上DSセンターの提供する教育プログラムの全体像を簡単に述べましたが、その中心となるのはやはり全学向けの正規授業科目です。特に、本学は私立総合大学として様々な学術領域を有しており、また学部・大学院生を合わせて5万人に迫る学生がいます。学生は入学方式も異なり、自身の興味や数学やプログラミング能力なども様々です。このようなバックグラウンドのもと、学生に向けていかに科目を構成し、これを運用するかについては唯一の正解はないと考えますが、最初の科目設置から数年経ち履修者数も増加傾向にあり、その成果も徐々に出てきているのではないかと考えております。「3.」以降でそれら正規授業科目のカリキュラムやそれに付随するサービスなどについて説明したいと思います。

3. データ科学教育科目

DSセンターは本学グローバルエデュケーションセンター（以下GEC）と協力してデータ科学授業コンテンツの開発を行っています。GECはDSセンター発足前より全学向けに基盤教育を展開しているセンターで、データ科学関連科目以外にも「アカデミック・ライティング」「数学」「情報」「英語」などのプログラムを提供しています。

正規授業科目のカリキュラムは、図2に示すようにA群からD群までの4つの科目群に分かれています。データ科学を初めて学ぶ学生は、まずA群を学ぶことを推奨しています。A群はデータ科学の基礎的な考え方と実践を学ぶ科目群で、B群はデータ科学を深く学ぶために必要となる数学やICT技術を学ぶ科目群です。ただし、B群科目は他の群の前提としてはおらず、より深く学びたい学生を対象としています。C群は自身の学ぶ専門

D群: エキスパートへ (Development)	自身の専門領域以外においてもデータ科学を活用できるための科目群
C群: 専門領域に活用 (Connected)	自身の持つ専門性や学術領域への接続科目群
A群: 考え方を身につける (Adequate)	データ科学の基礎的な考え方とその実践を一通り学ぶ科目群
B群: 基礎・周辺知識 (Basic)	より深く学び活用するために必要な数学やICT等の科目群

図2 データ科学科目群

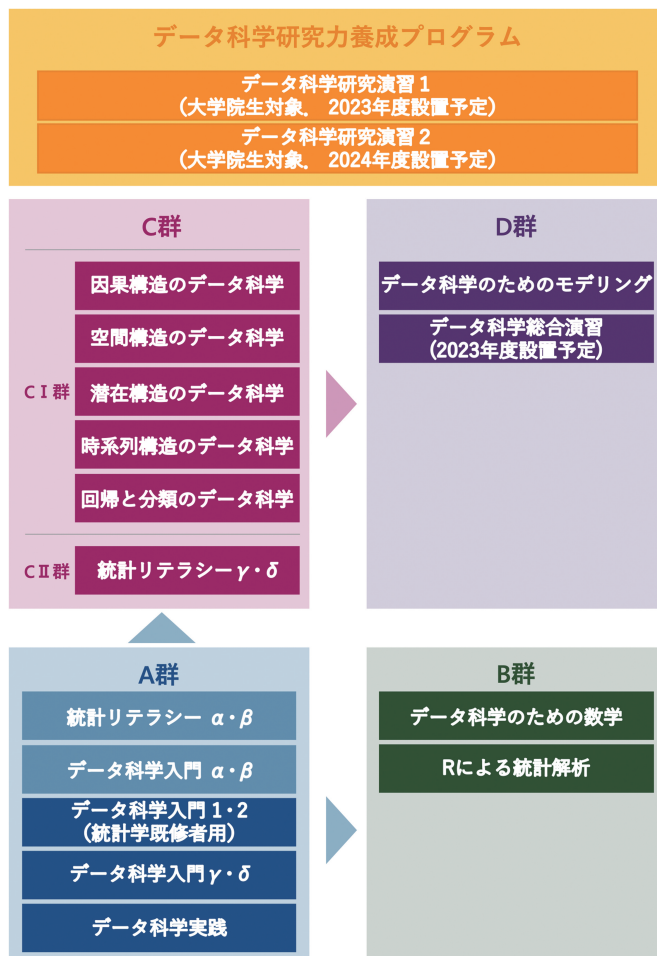


図3 データ科学教育プログラム科目

学術領域にデータ科学を活用することを目的とした科目群です。D群はC群までに培ったデータ科学の知見を一般化して、自身の専門学術領域以外にもデータ科学を適切に活用できるようになることを目的とした科目群になっています。

具体的な科目の構成は図3のとおりです。全てクォーター科目で提供しています。先に説明しましたように、初めてデータ科学を学ぶ学生にはまずA群科目を推奨しています。データ科学入門 $\alpha \sim \delta$ （以下DS入門シリーズと略す）は、データ科学の考え方の基礎を1年間で一通り学ぶ科目群です。特徴は統計学の基礎（記述統計、推測統計）や多変量解析、機械学習の内容を分けずに、統一的な考え方を通してデータ科学を学べるように配慮している点です。またDS入門シリーズでは、理論として学んだ内容をプログラミング言語Pythonの演習を通して実践し、その内容の理解と実践の両者を身につけることをもう一つの特徴としています。もう一方の統計リテラシー $\alpha \sim \delta$ は、オーソドックスな統計学の基礎的な内容を一通り

学ぶ科目群です。さらに、データ科学実践は、データ科学の基礎を一通り身につけた学生を対象にして、データ分析を行うプロセスを実際に手を動かして実践しながら学ぶ科目です。

C群の各科目の内容は科目名に表現されている通りで、対象となるデータの構造及び分析の目的に着目した観点で分かれています。これらのそれぞれの科目においても、統計学並びに機械学習を明確に分離することはせず、データ科学の統一的な視点から学ぶことができます。またA群科目と同様Pythonなどでの演習を同時に学ぶ内容となっています。D群科目は、C群科目を3科目以上学んでいることを前提として、それらを汎化してデータ科学を俯瞰するための「データ科学のためのモデリング」と、自身の専門でないデータに対して実際に分析を一から行い、分析目的を達成する方法を実践的に学ぶ「データ科学総合演習」があります。

現在さらにそれらを修めたあるいはそれに相当する知識のある修士学生向けにデータ科学研究力養成プログラムを2023年度後期より展開予定ですが、こちらの内容については紙面の都合上割愛いたします。DSセンターのホームページ^[4]などをご参照いただければと思います。

またここでの科目の一部は英語科目としても提供しており、英語化されていない科目も今後英語化予定となっています。

4. 本学データ科学認定制度

「3.」で述べましたとおり科目は全てクォーター科目ですが全部で20におよぶ科目を新規科目として準備してきました（いくつかは今後設置予定）。それらの科目群の中でどのように勉強を進めていけば良いか分からない学生に対して明確な目標を提示するために、独自のデータ科学認定制度（以下認定制度と略す）を設置しています。この認定制度では履修者のデータ科学に関する能力を保証する4つの級を設置し、級毎に到達目標を明示することで、各学生の興味関心に合わせたデータ科学の学習機会を提供しています。各級の定める要件を満たした学生に対しては証明書が発行されます。認定制度における4つの級とそれぞれの到達目標並びに取得条件は図4の通りです。図3、図4から分かるように、A群の前半の単位



図4 本学データ科学認定制度要件

取得がリテラシー級に対応しており、データ科学実践を含む後半まで単位を取得すると初級の認定を取得できます。ここで、「データ科学入門1・2（統計学既修者用）」とは、統計学を学んだことのある学生を対象として、DS入門シリーズの考え方を短くまとめて学ぶことができる科目です。このように、通常の統計学から学んでも初級までの認定を取得した学生はデータ科学の統一的な考え方を身につけることができていることを保証する制度としています。また中級と上級はそれぞれC群とD群に対応するように設計されており、データ科学を学ぶ自身の目的に照らし合わせて到

達目標を学生が設定し、学習のモチベーションを高めることに役立てることを期待しています。さらにはデータ科学関連のインターンシップなど産学連携などにおいても、認定制度の級の提示によりデータ科学の学びの到達状況を説明できるなどの利点も生まれています。このように、学生のキャリア教育の一端にも本認定制度が活用され始めている状況です。

5. 数理・データサイエンス・AI教育プログラム認定制度との関連

概要で述べましたように、本学データ科学認定制度のリテラシー級はMDASHのリテラシーレベルの認定を受けました。また同認定制度初級は同じくMDASHの応用基礎レベルプラスの認定を受けております。特に応用基礎レベルプラスの認定理由として「多くの学生にデータ科学教育プログラムを展開するためにフルオンデマンド科目として毎クォーター開講しているとともに、予約制のオンライン対面指導などで学生へのサポートが充実を図る取組みを実施している。」となっています。これはまさに本学の提供するデータ科学教育プログラムの特徴であり、5万人規模の学生にプログラムを提供するための施策の一つです。「6.」でこの特徴について述べていきたいと思えます。

6. データ科学教育プログラムの特徴

(1) 新規科目として設置

データ科学教育プログラムでは、「統計リテラシーシリーズ」4科目を除いた全ての科目を新規に設置しています。これにあたってまず「本学の学生に必要な内容は何か」「身につけてほしい力は何か」について様々な学部で常設の科目を調査・吟味しました。その上でDSセンター所属の教員で各科目内容を相談しながら検討、コンテンツを作成しています。特に従来統計学や多変量解析あるいは機械学習の分野で個別に発展してきた手法を「意思決定」の観点から整理し直していることが特徴の一つです。また数学的な内容についてはなるべく直感的に理解できるように図や例などを交えて説明するとともにPythonなどのプログラミングを利用することにより、分析の具体的なイメージが理解できるように配慮しています。これらにより初めての学生でも見通しよくデータ科

学について学ぶことができます。考え方について興味がある方はテキスト「データ科学入門IおよびII」を参照ください^{[2], [3]}。

(2) オンデマンド科目としての提供

全学生向けに科目を提供するために科目のほとんどをオンデマンド科目として提供しています。特にMDASHリテラシーレベル、応用基礎レベルに対応した科目群は全てオンデマンド科目です。この理由としては

- 本学は複数のキャンパスを有していること
- 学部で学ぶ専門科目が増える中で余裕を持って履修可能なこと

があげられ、これらの状況下で全学に向けて科目を提供するためにオンデマンド科目を採用しています。学生はいつでもどこでも空いている時間を見つけて科目を履修することが可能です。また各科目の1講義は3～4個のモジュールと呼ばれる個別の内容から構成されております。各モジュールは、主に「講義動画、小テスト、プログラミング」の三つからなり、各モジュールの講義動画は長くても20分程度です。100分間動画をずっと視聴するというのではなく、各モジュールを小まめに学んでいくことができます。この点も空いている時間の学修に配慮している点と言えます。

(3) 学生サポートサービスの充実

オンデマンド科目として設置する上で学生にとって不利益になり得る点が教員とのコミュニケーション、特に講義内の疑問点の解消があげられます。先に述べたように全学を対象にしており、数学やプログラミングに関する事前知識に差があるため、これらに特に配慮したいくつかの教育支援を行っています。

1) データ科学履修相談

DSセンターではデータ科学を学びたい学生からの相談を受け付け、その学びたい目的や内容に応じて履修科目や履修計画に関するアドバイスを行う教育支援を実施しています。この履修相談内容として、学生自身の将来の研究やキャリアも見据えてデータ科学を学びたいといった要望も多く、そのような場合に現時点でどのような学修をすべきかを知りたいケースが多々あります。本教



図5 データ科学履修相談

育支援では、このようにその学びたい目的に応じて相談・サポートを実施しています（図5）。

2) Learning Assistant(LA)による質問対応支援

学生にとって理解が難しい箇所や、演習におけるプログラミングで詰まってしまった場合にそれらの解消の手助けをするLA制度を採用しています。LAはデータ科学の十分な知識を有する大学院生（場合によっては学部生）で、データ科学に関する授業レベルの質問対応を行ってくれています。この質問対応方法にはさらに以下の形態があります。

① 本学3号館2階における対面指導

本学3号館2階には、授業期間中であればLAが常駐している部屋があります。質問のある学生はいつでもそこにいけば質問をすることができます。

② LMS掲示板への質問の書き込みとその返信

本教育プログラムはLMSを通じて提供していますが、LMS上の掲示板にて質問を受け付けています。本掲示板は受講生であれば誰でも閲覧・書き込みが可能なため受講生自身が他の受講生の質問に答えることも可能です。

③ チャットによる質問対応

授業期間中であればチャットを送ることでより3号館2階に常駐しているLAに質問をすることができます。他キャンパスに通学しており3号館を訪れる時間のとれない学生はこちらを利用することが多いです。

④ オンライン対面指導

チャットでの質問から必要と判断された場合には、Zoom（あるいはSkype）を利用したオンライン対面指導に移行することがあります。

学生はこのように用意されたいずれかのチャンネルを通して質問を行い、LAから回答を得ることが可能です。またLAと教員間ではこれら質問を共有しており、LAでは対応しきれない質問に対しては教員が回答することもあります。

また同様のサービスとして「データ科学研究相談」というサービスもありますが、こちらは卒業論文や学会発表などの研究レベルにデータ科学を活用する際に、DSセンターの教員のコンサルティングを受けることができるサービスです。

7. まとめ

本稿では本学の全学データ科学教育プログラムとして、特に正規授業科目のカリキュラムおよびその特徴について説明しました。本カリキュラムではデータ科学の統一的な考え方に軸を置き、また演習でPythonを用いて実践することによって、知識とスキルの両者を同時に学ぶ教育プログラムとなっています。これらの大きな目的の一つは、先に述べました自身の学術領域における専門性に、データ科学を活用できる人材の育成です。その意味で本科目の履修者数は年々増加してはおりますが、全体から見ますとまだまだ少ないと考えています。データ科学の重要性について積極的に広報していく必要性を感じています。

本プログラムの特徴の一つであるカリキュラムに対する支援についても述べました。具体的にはデータ科学履修相談及びLAによる質問対応により、学生の学修目的や知識レベルに応じて相談や質問をいつでも行える体制を構築しており、全学の多くの学生に対してきめ細かいデータ科学教育を提供しています。

より詳しく知りたい方はホームページなどもご覧いただければと思います¹⁾。

参考文献および関連URL

- [1] 早稲田大学データ科学センターホームページ <https://www.waseda.jp/inst/cds/>
- [2] 松嶋敏泰監修, データ科学入門I-データに基づく意思決定の基礎-, サイエンス社, 2022.
- [3] 松嶋敏泰監修, データ科学入門II-特徴記述・構造推定・予測 回帰と分類を例に-, サイエンス社, 2023.