

数理・データサイエンス・AI教育の紹介

千葉大学の数理・データサイエンス・AI教育 取組みの概要



千葉大学大学院理学研究院教授
データサイエンス教育実施本部副本部長 **松元 亮治**

1. はじめに

本学は首都圏にある国立大学で、国際教養学部、文学部、法政経学部、教育学部、理学部、工学部、園芸学部、医学部、薬学部、看護学部の10学部と大学院等から構成されています。学部学生の入学定員は学年あたり2,317名です。メインキャンパスである西千葉キャンパス（写真1）は東京駅から電車で45分という便利な場所にあります。これに加えて、亥鼻キャンパス（医学部、薬学部、看護学部）、松戸キャンパス（園芸学部）、柏の葉キャンパス、墨田キャンパスがあります。



写真1 本学西千葉キャンパス

本学では1994年度に教養部廃止と並行して全学生を対象とする情報処理教育の実施体制、演習設備を整備し、情報リテラシー教育を実施してきました。その実績を踏まえて2018年度に文部科学省「大学における数理・データサイエンス教育の全国展開」に「高大接続・学部・大学院に至るまで一貫した数理・データサイエンス教育とスマートラーニングによる全国・海外展開を含めた『千葉大学モデル』の構築」事業を申請し、協力校に選定されました。これを受けて、数理・デー

タサイエンス科目の全学必修化を検討し2020年度から数理・データサイエンス科目3単位を全学必修にしました。また、全学副専攻プログラム「数理・データサイエンス教育プログラム」を開始しました¹⁾。以下では、これらの取組みの概要を紹介します。

2. 数理・データサイエンス・AI教育の運営体制

本学では数理・データサイエンス・AI教育を全学体制のもとで実施しています。図1に原稿執筆時点での運営体制を示します。教育改革組織である国際未来教育基幹に「データサイエンス教育実施本部」を設置し、全学的な数理・データサイエンス・AI教育の舵取りをしています。

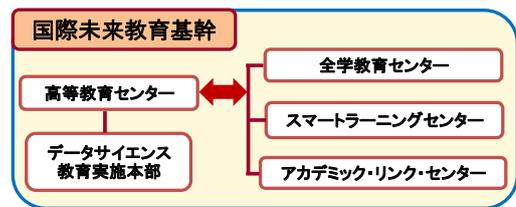


図1 本学における数理・データサイエンス・AI教育の運営体制

数理・データサイエンス科目の運営は様々な部局の約50名の常勤教員によって構成される「数理・データサイエンス教員集団」が担っています。この集団により、共通教材の開発、時間割作成、担当教員FD等が行われています。

協力校事業の当初計画にあった「スマートラーニング」については、新型コロナ感染拡大に伴い、オンデマンド教材作成等が一気に進み、2020年5月から開始された授業で活用されています。オンライン授業やハイフレックス授業はスマートラ

ーニングセンターの支援のもとでMoodleとGoogle Workspace等を用いて実施されています。

3. 数理・データサイエンス科目の位置づけ

本学の学士課程のカリキュラムを図2に示します。数理・データサイエンス科目は普遍教育科目の学術発展科目群に位置づけられています。なお、本学では教養教育のことを「普遍教育」と呼び、全学教育センターによって運営されます。

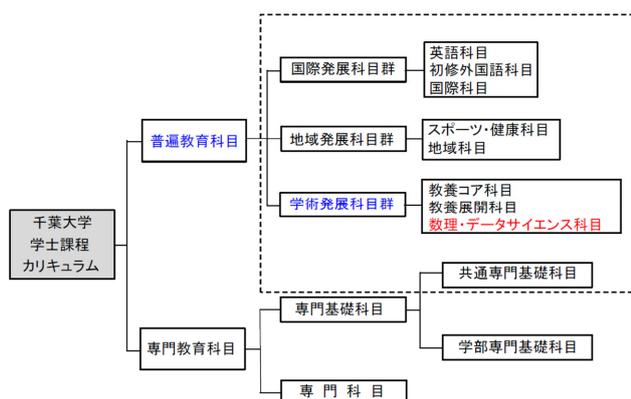


図2 本学の学士課程のカリキュラム

数理・データサイエンス科目には「基礎」と「展開」の科目があります。数理・データサイエンス科目（基礎）は、数理・データサイエンス・AIモデルカリキュラムのリテラシーレベルをカバーする2単位（90分×15回）の科目で、学部・学科別のクラス編成で実施されています。数理・データサイエンス科目（展開）は、統計基礎、プログラミング、実データ解析など数理・データサイエンス・AIモデルカリキュラムの応用基礎レベルまで含めた科目群で、学生は自分に合った科目を選択できます。1科目（1単位）必修となっています。

数理・データサイエンス科目3単位（基礎2単位、展開1単位）を必修とする教育プログラムは、2021年度に文部科学省「数理・データサイエンス・AI教育プログラム認定制度」の「数理・データサイエンス・AI教育プログラム リテラシーレベル（MDASH-Literacy）」に認定され、独自の工夫・特色を有する教育プログラムとして「数理・データサイエンス・AI教育プログラム リテラシーレベルプラス（MDASH-Literacy+）」にも選定されました。

4. 数理・データサイエンス科目（基礎）の内容

従来から実施されていた情報リテラシー科目を引き継ぐ科目で、全学共通カリキュラムによる講義パート（9-10週）と学部・学科の特性に応じた計算機を用いた演習パート（5-6週）から構成されます。講義パートは、リテラシーレベルモデルカリキュラムの「導入」、「心得」、「基礎」をカバーする以下の内容となっています。

（1）情報活用社会の到来

様々な機器がコンピュータネットワークを通して結ばれ、商品の購入情報、個人の位置情報などが瞬時に共有される情報社会の現状について学びます。また、取得されたデータを処理することによって、問題を解決したり、判断を行う材料となる情報が得られることを具体的な事例を通して理解させます。計算機の処理能力の向上によって、膨大なデータから有用な知見を得ることが可能になったことを示し、データを活用することによって成立する社会の到来が個人の生活とも密接に結びついていることを学びます。

（2）データと情報の表現

データには数値データ、文字データ、画像・映像データや音声データなどがあること、これらを計算機を用いて処理したり、情報ネットワークを通して送受信するためにデータがビット列として表現されることを学びます。2進数を用いた数値データの表現、文字コード、アナログ情報のデジタル化、画像のデジタル表現と情報圧縮について理解させます。遺伝情報を担うDNAの塩基配列データが活用されている事例、身近な情報表現であるバーコードやQRコードについても学びます。

（3）計算機と情報ネットワークの仕組み

計算機のハードウェアとソフトウェアの基礎を学びます。フォンノイマン型計算機の仕組み、主記憶装置と中央処理装置、問題を解決する手順であるアルゴリズムと、それを計算機向きに表現したプログラムについて講義します。また、情報社会の基盤となっているインターネットと、その上で利用できるサービスについて解説し、パケット交換、IPアドレス等を理解させます。

(4) 情報・データ倫理とセキュリティ

コンピュータウイルスなどのマルウェアやフィッシングメール等による被害、情報漏洩等の事例を紹介し、これらの情報セキュリティ上の脅威への対策について解説します。また、公開鍵暗号や電子署名など、セキュアな通信や個人認証の方法を理解させます。個人情報の保護に関する法律、EU一般データ保護規則、知的財産権とその保護についても講義します。データを扱う上での不正事項として、捏造、改ざん、盗用等があることを示し、レポート等を作成する際の参考文献の引用方法等についても学びます。

以上の各項目について本学で作成したオンラインテキストと動画教材が用意されており、Moodleを用いた視聴管理ができるようになっています。また、理解度を確認するための問題バンクが利用可能で、小テストなどに活用されています。独立行政法人情報処理推進機構（IPA）が公開している情報セキュリティに関する動画、情報処理技術者試験の過去問なども活用されています。

演習パートではデータ解析方法と解析結果の可視化について、計算機演習を行います。データの平均値や分散・標準偏差の計算方法、実験・観測データに含まれる誤差の扱い、標本の抽出方法、クロス集計などについて解説し、データの分布をヒストグラムとして表現したり、データ間の関係を散布図として表現する演習を行います。学部・学科に応じて、表計算ソフトを用いて演習を行うクラスと、Python言語を用いたデータ解析を行うクラスがあります。実データを用いた演習を通してデータを読み解く能力を高め、不適切に作成されたグラフや数字に騙されないデータリテラシーを養います。

演習クラスは原則として対象とする学生が所属する学部・学科の専任教員が担当しています。1クラスの学生数は40名～105名です。2022年度



写真2 総合校舎に設置された情報処理演習室

には31クラスが開講されています。演習設備としては総合校舎に100～105台の教育用端末を設置した演習室（写真2）が3室、統合情報センターにも演習室があります。新型コロナウイルス感染拡大に伴い、演習は対面型の授業と学生所有のPCを用いたオンデマンド方式を併用しています。なお、本学はマイクロソフト社との包括契約を結んでおり、学生は自宅からOffice365を利用できます。Python言語の演習にはGoogle Colaboratoryを利用しています。

5. 数理・データサイエンス科目（展開）の内容

数理・データサイエンス科目（展開）はリテラシーレベルモデルカリキュラムの「選択」と応用基礎レベルの内容を扱う科目で、2022年度にはデータサイエンスA（6クラス）、データサイエンスB（8クラス）、データサイエンスC（3クラス）、データサイエンスD（1クラス）に加え、「野球観戦に生きるデータ科学」、「Rによるアンケート調査の集計」、「中級データサイエンス」、「応用データ処理技術」等、合計26クラスが開講されています。各クラスの定員は50～130名、内容は以下の通りです。

(1) データサイエンスA

データ解析のための基本的な数理を学ぶ科目です。高等学校で数学Ⅲを履修していることを前提としない科目ですので、理工系以外の学生も履修することができます。最初に、1変量データの要約として、標本平均、標本中央値、標本分散、偏差値について学び、ヒストグラムや箱ひげ図について解説します。次に、2変量データの要約として標本共分散、標本相関係数、散布図を学び、最小2乗法による直線回帰について解説します。後半では確率分布と統計的推定を扱います。2項分布、正規分布などの基本的な確率分布を理解させ、母集団と標本、正規母集団の平均の区間推定（信頼区間）を学びます。

(2) データサイエンスB

データ解析やプログラミングに関心のある学生向けの科目です。データからの情報抽出とグラフによる可視化、データ間の関係を表す共分散、相

関係数の算出、最小2乗法による回帰直線などについて学びます。プログラミング演習はPythonまたはR (RStudio) を用いて行い、オープンデータの解析・可視化などの課題に取り組みます。後半には発展課題として機械学習によるデータの分類(クラスタリング)等を取り上げ、ニューラルネットワークを用いた深層学習についても解説します。講義内容の例が図3の小冊子にまとめられています。



図3 データサイエンスBの講義内容例
<https://mds.chiba-u.jp/files/pamphlet/coll/>

(3) データサイエンスC

情報科学の入門的な講義を行います。確率を用いて情報を定量化する方法と情報エントロピーについて解説した後、通信路を通して伝えることができる情報の大きさを求める方法について議論します。条件付き確率やベイズの定理、マルコフ過程など機械学習の基礎となる概念も扱います。情報圧縮、公開鍵暗号などについても解説します。

(4) データサイエンスD

コンピュータサイエンスの入門的な講義を行います。基本的な論理演算ができる部品を組み合わせることによって、計算や記憶が可能になることを示した後、電子計算機の仕組みについて解説します。具体的なプログラミング言語としてアセンブリ言語とPythonをとりあげ、これらを用いて基本的な手順(アルゴリズム)を表現する方法を学びます。ニューラルネットワークを用いた深層学習等、計算機の新たな可能性についても述べます。

(5) 野球観戦に生きるデータ科学

野球データを利用して、基礎的な統計知識を実践的に学習することで、統計学への理解をより深めることを目的とします。打率や防御率などの伝統的な指標から、セイバーメトリクスと称される近年生まれた指標も紹介します。それらの有用性を相関分析などで検証する作業を通して、データの扱い方や各種検定方法を学習します。また、複数の指標を用いて重回帰式を作成して、戦術や戦略面での有効性(未来予測)について考えていきます。データ収集方法の学習や、スタジアムで観戦しながら試合予測をする実地調査も実施します。

(6) Rによるアンケート調査の集計

Rを用いて、アンケート調査の集計に取り組みます。これを通してRに慣れ親しむとともに、Rでの基本的なデータ処理や統計処理について学習します。より具体的には、大学で行われている学生調査などを対象に、Rを用いて集計を行うことで、調査集計についての考え方とデータの前処理(データハンドリング)や、データの可視化、レポート生成の技術について扱います。

(7) 中級データサイエンス

統計・検定について講義します。データサイエンスAとは異なり、高等学校において数学Ⅲを履修していることを履修要件とします。また、プログラミングの演習も行います。

(8) 応用データ処理技術

音声、画像・映像、主観的データなどのデータの扱い方を学ぶ科目です。具体的には、連続信号のデジタル化の原理と手法、メディアごとに異なるデータ表現や特徴抽出手法、教師なし学習と教師あり学習、主観的データの取得方法や分析手法、仮説検定などを扱います。プログラミングにはPythonを用い、Google Colaboratoryを用いて各種プログラムを作成することを通じて中級レベルのスキルを習得します。

以上の科目の共通教材として、本学においてデータサイエンスを活用した研究を実施中の研究者による研究紹介動画を作成しました。一部は次ページ図4のように公開されています。



図4 データサイエンスを活用した研究紹介動画
<https://mcs.chiba-u.jp/materials.html>

6. 数理・データサイエンス副専攻の設置

本学では、2020年度からの入学生を対象として、全学副専攻プログラム「数理・データサイエンス教育プログラム」を開始しました。このプログラムは「国際日本学」、「ローカル・イノベーション学」につづく3番目の副専攻プログラムです。

数理・データサイエンス副専攻プログラムは、数理・データサイエンスに関する基礎的な教養をベースに、各学部における専門的な数理・データサイエンスを極めることにより、数理・データサイエンスに係る知識を活用し、社会の問題を解決できる人材を育成することを目的としています(図5)。

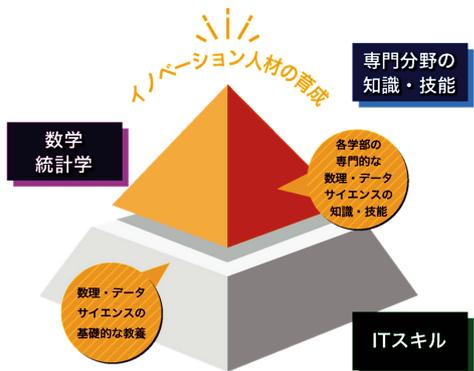


図5 全学副専攻プログラム「数理・データサイエンス教育プログラム」の目的

「数理・データサイエンス教育プログラム」では、表1に示す取得要件に定められた所定の単位を修得することにより、「修了証書(30単位)」または「履修証明書(20単位)」を卒業時に発行

表1 「数理・データサイエンス教育プログラム」の取得要件表

修了証書の取得要件表		必要単位数
科目区分		
普遍教育科目	学術発展科目群 (数理・データサイエンス科目(基礎))	2
	学術発展科目群 (数理・データサイエンス科目(展開))	1
	数理・データサイエンス科目及び 教養展開科目「データを科学する」	1~3
	計	4~6
共通専門基礎科目	数学・統計学(微積分学)	4
	数学・統計学(線形代数学)	4
	数学・統計学(統計学)	4
計	12	
専門教育科目 自由選択	数理・データサイエンスに 関する専門内容	12~14
	合計	計30単位

履修証明書の取得要件表

履修証明書の取得要件表		必要単位数
科目区分		
普遍教育科目	学術発展科目群 (数理・データサイエンス科目(基礎))	2
	学術発展科目群 (数理・データサイエンス科目(展開))	1
	数理・データサイエンス科目及び 教養展開科目「データを科学する」	1~3
	計	4~6
共通専門基礎科目	数学・統計学(微積分学)	2
	数学・統計学(線形代数学)	2
	数学・統計学(統計学)	2
計	6	
専門教育科目 自由選択	数理・データサイエンスに 関する専門内容	8~10
	合計	計20単位

します。この副専攻プログラムを履修するには、定められた期間中に履修登録を行い、計画的に履修を進めます。

修了証書を取得するには数理・データサイエンス科目(基礎、展開)3単位に加えて、数理・データサイエンス科目または教養展開科目「データを科学する」から1-3単位、共通専門基礎科目の微積分学、線形代数学、統計学、合計12単位、専門教育科目における指定科目から12-14単位を修得する必要があります。履修証明書を取得するのに必要な普遍教育科目の単位数は修了証書と同じですが、数学は微積分学、線形代数学、統計学6単位、専門教育科目の指定科目は8-10単位が要件となっています。理学部・工学部の学生は修了証書のみ取得可能です。

教養展開科目「データを科学する」には、情報

セキュリティ（入門、実践）、地球環境とリモートセンシング、デジタルクリエイティブ基礎などの科目があります。専門教育科目の指定科目には、すべての学部の学生が履修できる全学共通科目（データ解析基礎論、知能システム入門、計算物理学、緑地環境情報学など28科目）と所属学部の学生のみが履修できる科目（国際教養学部 6科目、文学部 33科目、法政経学部 15科目、教育学部 2科目、理学部 27科目、工学部 57科目、園芸学部 9科目、医学部 2科目、薬学部 1科目、看護学部 4科目）があります。

7. 大学院共通教育科目

本学では大学院生に対して、①自律的・自立的・組織的に研究を行うにあたって基盤となる資質・能力、②社会経済の各分野において指導的役割を果たすとともに、国際的にも活躍できる高度な専門的能力、③自身の専門的知識・技能・経験を他者にわかりやすく伝え、他者や組織の成長を促す能力を高めることを目的として、大学院共通教育科目を開講しています^[2]。

2022年度には、数理・データサイエンス・AI関連の共通スキル科目として「データサイエンス」、「データ科学プログラミング」、「機械学習実践」、「デジタル・ヒューマニティーズ入門」が全学の大学院生を対象として開講されています。「データサイエンス」ではRを用いたデータ解析、「データ科学プログラミング」ではPythonを用いたデータ解析、「機械学習実践」では深層学習を中心とする機械学習の講義と演習、「デジタル・ヒューマニティーズ入門」では人文社会科学系の学問分野におけるデータサイエンスの応用としてのデジタル・ヒューマニティーズに関する研究動向を、特に海外での先進的な事例を中心に学ぶことを目的としています。

8. 高大接続事業

本学は1998年度から、高校2年生から大学に飛び入学することを可能にした「先進科学プログラム」を実施してきました。情報工学コースでは日本情報オリンピックの予選に出場していることが受験資格になっています。

2020年度からは科学技術振興機構（JST）グローバルサイエンスキャンパス事業「Society 5.0を

創出する未来リーディング人材育成 ASCENT Program」を実施中です^[3]。このプログラムでは高校生を対象として科学技術の基礎力に加えデータサイエンスの素養を身につけ、大学の環境を生かして研究を行う機会を提供しています。選抜の上、受講している高校生は約40名です。

9. 産業界との連携

大学生向けの数理・データサイエンス・AI教育教材について、ベネッセ・コーポレーションとの共同研究協定のもとに、教材内容、LMS等について検討しました。また、2020年度には開発された教材を数理・データサイエンス科目（展開）の授業で試用し、補助教材としての有用性等を検証しました。さらに、アドビ社と提携して「デジタルクリエイティブ基礎」というデータの可視化を中心とする科目を創造性や問題解決能力を涵養する教養教育科目として共同実施しています。これに加え、帝国データバンクの協力のもと、カリキュラムの体系性・実践性についての検証を進めています。

10. 今後の展開

本学の数理・データサイエンス・AI教育の概要について紹介しました。今後は、全学生を対象とするプログラムと数理・データサイエンス・AIのエキスパートを育成するプログラムの間をつなぐプログラムを加えることを予定しています。また、運営体制についても見直しを進めています。

本学では医学研究院に治療学人工知能（AI）センターが設置される等、数理・データサイエンス・AIを活用した研究が大きく進展しつつあります。研究と教育を両輪として、高大接続、リテラシーレベルから大学院における研究者・高度技術者育成まで一貫した数理・データサイエンス・AI人材育成を進めていく計画です。

関連URL

- [1] 千葉大学数理・データサイエンス教育プログラム
<https://mds.chiba-u.jp/>
- [2] 千葉大学大学院共通教育
<https://www.cphe.chiba-u.jp/graduate-common/>
- [3] JSTグローバルサイエンスキャンパス「Society 5.0を創出する未来リーディング人材育成 ASCENT Program」
<https://gsc.e.chiba-u.jp/program/>