

大規模私立総合大学におけるオンデマンドデータサイエンス教育の取組

岡部格明*1・宿久洋*1
Email: mokab.0328@gmail.com

*1: 同志社大学

◎Key Words 数理・データサイエンス・AI教育,大規模授業,オンライン学習,オンラインアセスメント

1. はじめに

近年、情報技術の発達により、大量のデータを収集、解析することが可能となった。これにより、AI やデータサイエンスに対する社会の期待が高まっている。AI・データサイエンスの利活用も広がりつつあるが、依然として、数理・データサイエンス・AI を身につけた人材が不足していることが指摘されている。2019年に政府は、AI 戦略2019⁶⁾をまとめた。この中で、AI、数理、データサイエンスの具体的な教育目標として、「文理を問わず、全ての大学・高専生(約50万人卒年)が、過程にて初級レベルの数理・データサイエンス・AIを習得することが、挙げられている。

これを受けた大学教育における取り組みとして、文部科学省は、数理・データサイエンス・AIの教育を進める施策として、2021年度から数理・データサイエンス・AI教育プログラム認定制度を開始した⁹⁾。2021年度からはリテラシーレベルの認定、2022年度からは応用基礎レベルの認定が開始されている。2021年度には、リテラシーレベルのプログラムとして、59大学、2短期大学、6高等専門学校計67の組織が実施しているプログラムが認定された。2022年度には、リテラシーレベルのプログラムとして、96大学、9短期大学、34高等専門学校の計139の組織が実施しているプログラムが認定され、応用基礎レベルのプログラムとして、大学等の単位で合計27の組織が実施しているプログラム、学部・学科の単位で合計41の組織が実施しているプログラムが認定された。

2. 同志社大学の取り組み

本報告では、同志社大学におけるAI・データサイエンス教育プログラムの取り組みおよび、その必修科目についての取り組みとその課題について紹介する。

同志社大学でも2022年度より、文部科学省の「数理・データサイエンス・AI教育」に関する認定制度に準拠した、同志社データサイエンス・AI教育プログラムを開始している⁸⁾。このプログラムはDDASH (Doshisha Approved Program for Data Science and AI Smart Higher Education) プログラムと呼ばれており、このプログラムの概要を図1に示す。同志社大学は、全14学部から構成される総合大学であり、全学で学生数が約2.5万人に及ぶ大規模私立大学である。DDASHプログラムの特徴として、特定の学部・に所属する学生だけではなく、全学の学生がプログラムを受講可能という特徴がある。

DDASHプログラムのアセスメント方法の1つとして、一般財団法人統計質保証推進協会が実施している「統計検定」の受験を推進している。同志社大学は、「統計教育

連携ネットワーク (JINSE)⁷⁾に参画しており、全ての学生は割引価格で受験可能である。DDASHプログラムはDDASH-L (リテラシーレベル)、DDASH-A (応用基礎レベル)、DDASH 副専攻、DDASH-E (エキスパートレベル)の4段階で構成されている。DDASH-L、DDASH-A、DDASH 副専攻は学部レベルのプログラムであり、DDASH-E は大学院レベルのプログラムである。

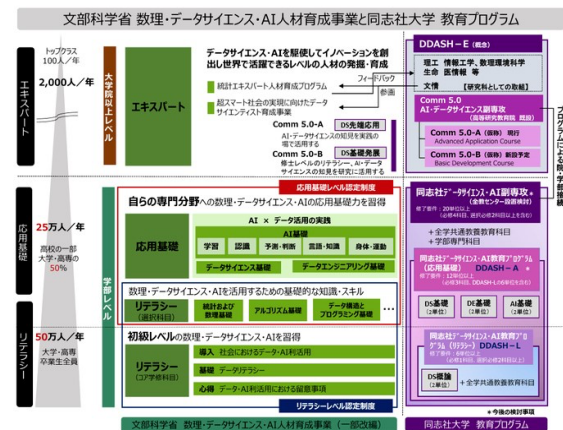


図 1: DDASH プログラムにおける「データサイエンス概論」の位置付け

リテラシーレベルに対応する DDASH-L は、実社会で目にするデータを適切に読み解き、使い方を判断できる水準の数理・データサイエンス・AIに関する能力取得を目的としたプログラムである⁸⁾。この教育プログラムは、必修科目である「データサイエンス概論」(2単位)の単位取得に加えて、プログラム選択科目として、全学教養科目として開講されている科目、「数学1」、「数学2」、「データサイエンス1」、「データサイエンス2」、「論理的思考の基礎(1)」、「論理的思考の基礎(2)」、「Statistics for the Social Sciences and Humanities」(各2単位)の中から4単位以上取得することによって、修了することができる。DDASH-L 修了レベルでは統計検定データサイエンス発展の合格レベルを目指す¹⁰⁾。

応用基礎レベルに対応する DDASH-A は、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や、AIを活用し課題解決につなげる機能力を修得し、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獲得を目的としたプログラムである⁸⁾。この教育プログラムは DDASH-L の修了要件に加えて、必修科目として「データサイエンス

基礎」,「データエンジニアリング基礎」,「AI基礎」(各2単位)の3科目の単位を取得することによって,修了することができる。DDASH-A 修了レベルで統計検定データサイエンスエキスパートの合格レベルを目指す⁽¹⁰⁾。

さらに拡充したプログラムである DDASH 副専攻は,数理・データサイエンス・AI教育(応用基礎レベル)の学びに加えて,各学部で開設している学部専門科目の履修を通して,エキスパートレベルである専門教育への学びにつなげることを目的としたプログラムである⁽⁹⁾。この教育プログラムは DDASH-A の修了要件に加えて,各学部の開講科目など対象とされている選択科目を合計20単位以上取得することによって,修了することができる。

3. データサイエンス概論

3.1 講義の概要

データサイエンス概論は, DDASH プログラム認定に際して, 2022 年度から開講されているオンデマンド講義である。この科目は, 同志社大学全学の学生が受講することができる。この講義は, 最終評価も含めてオンラインで行うという形のフルオンデマンド形式で開講されている。この講義の内容は, 数理・データサイエンス教育強化拠点コンソーシアム⁽⁶⁾において策定された, リテラシーレベルに含まれる,「導入」,「基礎」,「心得」の項目に準拠した内容となっている。この講義は, 全15回から構成されており, 講義は4人の教員と6名の外部講師によって行われる。講義内容は, 大きく「ガイダンスと導入」,「社会におけるデータ・AI の利活用」,「データリテラシー」,「データ・AI 利活用における留意事項」の4つの項目から構成される。

「ガイダンスと導入」においては, モデルカリキュラムにおける「社会で起きている変化」や「データ・AI 利活用のための技術」を取り扱う。「社会におけるデータ・AI の利活用」は, 外部講師6名による講義であり,「データ・AI の利活用領域」,「データ・AI 利活用の現場」,「データ・AI 利活用の最新動向」を取り扱う。調査系企業や公官庁, 製造系企業, 製薬系企業, IT系企業など, 様々な分野の外部講師(総務省統計局・独立行政法人統計センター, 株式会社日経リサーチ, アステラス製薬株式会社, 株式会社リクルート, 京セラ株式会社, 株式会社 Agoop)を招聘することがこの講義の特色の1つとなっている。

「データリテラシー」では, モデルカリキュラムの「基礎」部分にあたる「データを読む」,「データを説明する」,「データを扱う」にあたる部分を取り扱う。具体的には,「様々なデータ」,「データの収集」,「1変数データの記述と要約」,「2変数データの記述と要約」,「データの視覚化」の5つのトピックを取り扱う。「様々なデータ」では, 種々のデータを尺度水準などいくつかの切り口で分類してその種類について説明する。「データの収集」では, データ収集の目的と方法について, 概念としての母集団と標本や実験研究と観察研究について説明する。「1変数データの記述と要約」では, 「中心」を表す指標, 「散らばり」を表す指標などのほかに, ローレンツ曲線とジニ係数, 価格指数を取り扱う。「2変数データの記述と要約」では, 2つの量的変数同士の関係性の記述, 2つの質的変数同士の関係性の記述を取り扱う。2つの量的変数同士の関係性の中では, 相関係数や偏相関係数, 擬似相関を扱う。2つの

質的変数同士の関係性の中では, 一般の分割表に関する関連性の指標, 2標本における分割表の関連性の尺度, 予測結果解釈のための分割表の尺度を取り扱う。「データの視覚化」では, 変数の数とデータの種類ごとに様々な視覚化の方法を説明する。「基礎」の部分は全8回の講義を用いて, 先述の内容を丁寧に教える。確率や推測統計についてはこの講義内では踏み込まずに, データの要約やデータの視覚化の内容を教えることにとどまっており, 記述的な事項を丁寧に教えることに重点が置かれている。

「データ・AI 利活用における留意事項」では, モデルカリキュラムの「心得」にあたる部分を取り扱う。具体的には「データ分析での注意」,「法の世界とデータ利活用」,「情報倫理, 情報セキュリティ」の3つのトピックを扱う。「データ分析での注意」では, 統計を専門とする教員が講義を行い, ここでは, 恣意的な視覚化, 因果関係と相関関係, これまでの歴史におけるデータ分析の失敗例を紹介する。「法の世界とデータ利活用」では, 法を専門とする教員が講義を行い, データ利活用やデータ利活用やAI活用によって個人のプライバシーが侵害される危険や市場の公正さが損なわれる危険に関する説明を行う。「情報倫理, 情報セキュリティ」では, 情報を専門とする教員が講義を行い, 情報を扱う上でのモラルやマナー, 関連する法律や情報の安全性に関する説明を行う。

3.2 実施上の工夫

講義実施上の課題として, 受講者数の多さと受講生の評価に関する難しさが挙げられる。この講義は全学からの受講を受け付けるため, 学生, 教員とも時間の確保が困難となる。特に, 本講義では講義期間中の15回の講義を複数教員で担当するため, 開講時間の柔軟さが求められる。開講時間の問題に加えて, 受講者数が多いことにより, 各回の細かな採点が困難になることが問題として挙げられる。そのことにより, 学生の取り組み状況の評価が困難となってしまう。

受講者の成績評価基準として, 次の得点構成を用いている。講義動画の視聴に対する平常点20%, 講義内容に関わる演習問題である講義内課題の提出状況を20%, 各回の確認テストとしての講義外課題の得点を30%, 15回の講義終了後の実施される確認テストとしての最終評価の得点30%としている。

各回の講義内課題は, 次週に解答例を公開しているため, 受講者の自主的な復習を期待したものである。各回の講義外課題は, LMS 上で選択問題, 数値入力問題として出題している。いずれの課題も講義公開週の土曜日に質疑応答の時間を設けているため, 受講者は疑問点をその時間に質問することができる。

最終評価は, 制限時間を60分とした確認テストを行う。出題範囲は各回の講義外課題として, 春学期講義では114問, 秋学期講義では122問の問題セットからランダムに20題出題した。これにより, 他の受講生とは異なる問題が出題される。この確認テストをLMS 上で各受講者に解答してもらう。受験可能回数が1回のみであり, 一度開始したら終了できないため, 操作方法などは事前の回答練習ページで練習してもらう。最終評価受験時には, 問題に関する質問や, トラブル発生時の対応ができないため, 質問やトラブルに関する問い合わせについては,

別途質問のためのフォームを設けることにより対応している(図2)。実際、受講生によっては、自身の受験環境などに起因するネットワークのトラブルなどが発生していた。これらの場合には、最終評価の受験可能期間終了後に別途対応を行なった。

成績の評価方法(事前に受講生に提示したスライド)

- 最終評価に対して質問がある場合
 - e-classの「最終評価に関する質問窓口」から入力を行ってください。
 - 入力する場合は最後まで試験を終えてから
 - 途中でテストを中断した場合、再開することができません
 - e-classのシステム上の問題と思われる場合
 - 「最終評価に関する質問窓口」の設問1でその旨を回答してください。
 - 「最終評価に関する質問窓口」の設問2で症状やエラーメッセージ、受験環境を記入してください。
 - 個別評価の問題に関する質問の場合
 - 「最終評価に関する質問窓口」の設問3で該当する問題番号と質問内容を記入してください。
 - このとき、どのようにその問題を考えたのかを記入してください。
- 質問内容への個別対応は、e-class側のシステムトラブルなどを除き原則として行いません。



DDASH

図2: 最終評価におけるトラブル発生時の窓口

3.3 受講者の状況

2022年度は春学期、秋学期通年で1657人が受講し、合格率は77.7%であった。各学期の内訳として、春学期は、939人が受講し80.6%が合格している。秋学期は、718人が受講し74.0%が合格している。春学期と秋学期ともに全14学部全ての学部からの受講生が見られ、春学期と秋学期に両方とも受講した受講生は見られなかった。

春学期登録者数(文理別)

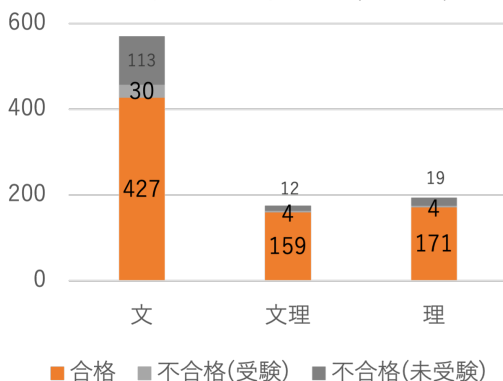


図3: 2022年度春学期の文理別登録者数

秋学期登録者数(文理別)

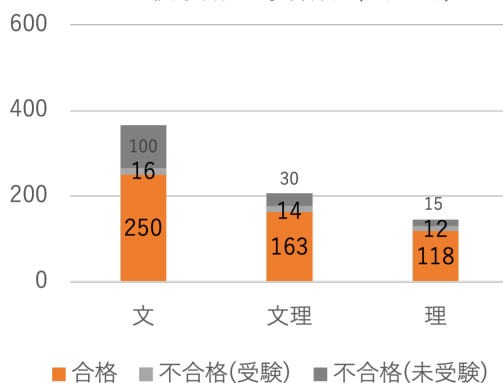


図4: 2022年度秋学期の文理別登録者数

春学期と秋学期の講義における受講者の文理区分および合否をそれぞれ図3、図4に示す。ここでの文理区分は、一般選抜入学試験の試験教科の文系型・理系型によって分けている。文系は、神学部、文学部、法学部、経済学部、商学部、政策学部、社会学部、グローバル・コミュニケーション学部、グローバル地域文化学部、国際教育インスティテュートの9学部、1組織である。国際教育インスティテュートに所属している学生は、文系6学部のいずれかに所属している。文理系は、文化情報学部、スポーツ健康科学部、心理学部の3学部である。理系は、理工学部、生命医科学部の2学部である。全体として、文系の受講生が多いことがわかる。春学期の文系学生の合格率は74.9%、文理系学生の合格率は90.9%、理系学生の合格率が88.1%であった。秋学期の文系学生の合格率は68.3%、文理系学生の合格率は78.7%、理系学生の合格率が81.4%であった。一方で、不合格者に限って見ると、不合格者全体のうち、そもそも最終評価を受験していないが一定数見られたことがわかる。春学期の文系学生で79%、文理系学生で75%、理系学生で82.6%、全体では79.1%見られた。秋学期の文系学生で86.2%、文理系学生で68.2%、理系学生で55.6%、全体では77.5%見られた。

春学期動画視聴割合

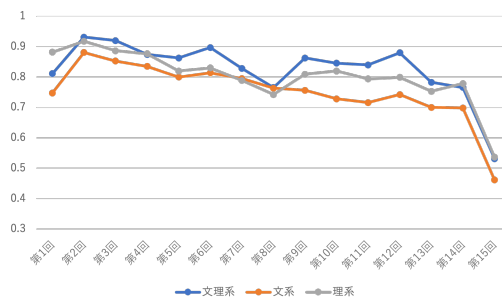


図5: 2022年度春学期の各回の講義動画視聴状況

秋学期動画視聴割合

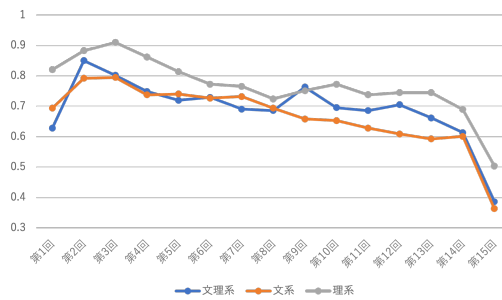


図6: 2022年度秋学期の各回の講義動画視聴状況

春学期と秋学期の講義における、文理別の動画視聴状況の推移を図5、図6に示す。各回の動画が概ね60分程度であるため、倍速視聴している受講生がいることを考慮し、動画公開期間中に動画ページに合計で20分以上の滞留している場合に動画を視聴していると定義した。春学期においても秋学期においても、全受講者に占める動画視聴者の割合は、講義が進むにつれて低下していることがわかる。文系受講者、文理系受講者、理系受講者のいずれも第6回から第8回で下がっている。しかし、第9回で文系受講者以外は再び上昇していることがわ

かる。第9回「データリテラシー(6) 2変数データの記述と要約(2)」では、2変数の質的データを取り扱う。文理系受講者及び理系受講者に着目すると、第6回から第8回の内容については、高校などですでに履修済みであることから視聴を飛ばしていることが考えられる。そのため、第9回は新たな内容であるため、視聴割合が再び上昇したことが考えられる。一方で、文系受講者については、単調に減少しているため、何らかの対策が必要であることが考えられる。

最後に、春学期と秋学期の講義における講義外課題の平均得点率をそれぞれ図7、図8に示す。これより、春学期、秋学期ともに第7回で大幅に得点率が下がっていることがわかる。第7回「1変数データの記述と要約(2)」では、1変数データのばらつきに関する問題と重み付き平均、データの標準化に関する問題を出題しており、特に重み付き平均の問題に関する正答率が低かった。

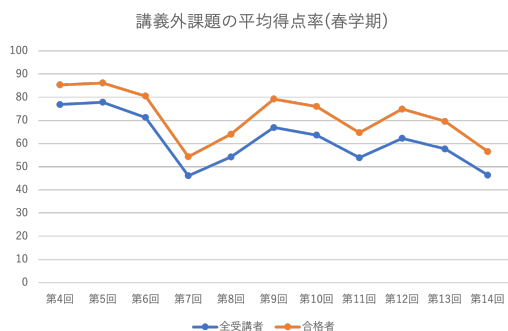


図7: 2022年度春学期の各回の講義外課題における平均得点率

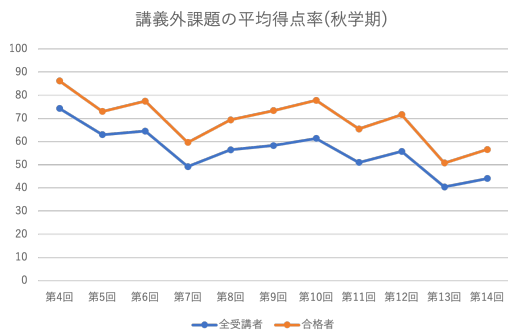


図8: 2022年度秋学期の各回の講義外課題における平均得点率

3.4 今後の方向性

この講義の今後の方向性に関して、講義の内容に関する改善、受講対象者の拡大、最終評価方法の見直しの3つの方向性がある。

講義内容に関して、第6回から第9回の内容や演習問題の改善が必要であると考えられる。高校で履修しているなど、ある程度わかっている文理系や理系の学生にとってはこの内容は簡単と思われる内容であり、そうでない文系の学生にとっては難しいと思われる内容になってしまっている可能性がある。そのため、文系の学生をフォローしつつ理系の学生を退屈させない内容にしていく必要があると考えられる。

受講対象者の拡大に関しては、2023年度の秋学期より

学校法人同志社の学内高校の生徒に対して、科目等履修生として、データサイエンス概論の履修を認める方向で進めている。特に第6回から第9回の内容については、高校生の受講に際して、検討を進めていく必要があると考えられる。

最後に、最終評価方法の見直しについて、現状のLMSシステムでは、悪意を持った受験者に対する対策を行うことができない。例えば、受験する際に生成AIを使うことや替え玉受験が可能になってしまう。本講義の扱う範囲のAI技術が含まれており、Open AIのChat GPTなどの生成AIについても重要な問題となっている。講義の中では、自身の学習を補助する型での利用を推奨しているが、必ずしもそのような使い方に留まらないことも考えられる。受講生の能力の保証を行う際には、生成AIや替え玉受験の対策を行うことが求められている。この問題への対処として、受験時のカメラ監視を含んだ本人認証、操作中のタブ移動やコピーの禁止、デスクトップ監視を機能としてもオンライン試験サービスの導入を検討している⁽¹⁾⁽²⁾⁽³⁾⁽⁴⁾。

4. おわりに

本稿では、大規模私立総合大学である同志社大学におけるオンデマンド開講講義、データサイエンス概論について説明した。この講義は最終評価も含めて、フルオンデマンドで開講している講義のため、受講生の評価が問題となるがLMSの活用によって、これらの問題に取り組んでいる。講義の内容については、高校で履修済みである部分に関して動画の視聴状況に変化が見られることから、今後対策が必要となっていることが示唆されている。講義及び最終評価で出題した問題のうち、正答率が特に低かった問題及びその詳細については、当日発表する。

参考文献

- (1) 株式会社EduLab. 『Check Point Z』. <https://checkpointz.com/jp/>.
- (2) 株式会社サーティファイ. 『スマート入試』. <https://smarte.jp/>.
- (3) 株式会社シェアウィズ. 『WisdomBase』. <https://wisdombase.share-wis.com/>.
- (4) 株式会社プロシーズ. 『Testable』. <https://www.proseeds.com/testable/>.
- (5) 首相官邸政策会議 (2019). 『AI戦略2019』. https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistrategy2019_fu_sanko.pdf.
- (6) 数理・データサイエンス教育強化拠点コンソーシアム (2020). 『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』. http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf.
- (7) 統計教育連携ネットワーク (2022). 『JINSE 版統計検定』. <https://jinse.jp/kentei>.
- (8) 同志社大学 (2023). 『「同志社データサイエンス・AI教育プログラム」(DDASH)』. <https://cgile.doshisha.ac.jp/ddash/overview.html>.
- (9) 文部科学省 (2022). 『「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・応用基礎レベル)」の認定・選定結果について』. https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/1413155_00011.htm.
- (10) 統計検定 (2021). 『統計検定 検定種別』. <https://www.toukei-kentei.jp/exam/>.