

学習意欲の予測モデル構築の経験に基づく 学校教育における教育 AI の利活用に向けた考察

小原 百々雅^{1,a)} 梶浦 照乃¹ 田村 みゆ² 秋信 有花² 倉光 君郎^{1,b)}

概要：

近年、データサイエンスや人工知能は、これからの社会活動を活性化させる基本情報処理として、その利活用が期待されている。この期待は、学校教育の現場においても例外ではない。しかし、データサイエンスや人工知能に関する情報は増えているが、学校教育に適用した事例は依然として少なく、ビジネス分野や医療分野で先行した事例を参考に手探りな状況である。本稿では、コロナ禍の下、オンライン・プログラミング演習で得られたデータに対し、機械学習アルゴリズムを適用した経験を事例として、今後の教育 AI の利活用を活性化させる方策を論じてゆきたい。

1. はじめに

近年、データサイエンスや人工知能はこれからの社会活動を活性化させる基本情報処理として、その利活用が期待されている。この期待は、学校教育の現場においても例外ではない。しかし、データサイエンスや人工知能に関する情報は増えているが、学校教育に適用した事例は依然として少なく、ビジネス分野や医療分野で先行した事例を参考に手探りな運用を行っている状況である。

本稿では、遠隔講義で得られたデータに対し、機械学習アルゴリズムを適用した経験を事例として、今後の教育 AI の利活用を活性化させる方策を論じてゆきたい。

本稿で紹介する事例は、オンライン・プログラミング演習における学習意欲の推定の試みである。2020 年度、コロナウィルス感染症の対策として、突然に始まったオンライン講義であるが、学生の理解度や意欲を把握しにくかった。そこで、我々は収集可能なデータ(自由感想文と学習ログ)から学習意欲を予測する機械学習モデルを作成し、学習意欲の低下兆候を早期に発見することを目指した。我々は、この目的のため、古典的な回帰分析から自然言語処理の最先端モデルである BERT までを用い、様々なデータサイエ

ンス技法を用いて予測モデルの構築を試みた。なお、2021 年度も引き続きオンライン講義としてデータ収集を行い、新たな分析を試みている。本稿は、どちらかといえば中間報告として、教育現場における機械学習を導入する点の気づきや課題、今後の展望を述べたい。2 節では、まずデータサイエンスと機械学習を外観する。3 節では、我々が取り組んだ学習意欲の推定を事例としたデータサイエンスと機械学習の適用事例を紹介する。4 節では、適用事例をもとに教育分野における人工知能活用のあり方を論じる。5 節では、本稿を総括し、今後の展望をまとめる。

2. データサイエンスと機械学習

本節では、本稿を理解するための予備知識をまとめながら、データサイエンスと機械学習を外観する。

2.1 データサイエンス

データサイエンスとは、データを用いて新たな科学および社会に有益な知見を引き出そうとするアプローチである。データを扱う手法として、統計学やコンピュータ科学を横断的に用いる。大別すると、次の2つが重要視される。

- データ分析: 統計学や可視化を用いて、得られたデータに対する説明力
- データ予想: 機械学習を用いて、未来に得られるデータに対する予測力

どちらの視点もデータサイエンスにとって重要である。本稿は、(著者らの専門性のため) 統計学よりもコンピュータ科学よりの視点に立ち、機械学習を導入した教育分野への活用を想定する。この想定は、教育 AI の一般的なイメー

¹ 日本女子大学理学部数物数学科
Department of Mathematical and Physical Sciences, Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan

² 日本女子大学大学院理学研究科数理・物性構造科学専攻
Graduate School of Science Division of Mathematical and Physical Sciences, Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan

a) m1816019om@ug.jwu.ac.jp

b) kuramitsuk@fc.jwu.ac.jp

ジに近いものと思われる。

2.2 機械学習

機械学習は、コンピュータビジョン（画像認識）、音声認識、自然言語理解の研究分野で長らく研究されてきたアルゴリズムである。共通するのは、データから予測モデルを構築し、その予測モデルから未知のデータに対する予測を行う点である。アルゴリズム自体は、データに対し高い普遍性を持ち、画像や音声などの領域を超えて予測モデルの構築に活用できる。

機械学習アルゴリズムは、教師なし学習と教師あり学習に大別することができる。特に、教師あり学習は、教育者の経験や知識をデータとして活用しやすく、教育 AI 実現の重要な手法となる。以下、少し詳しく解説をしてみる。

教師あり学習は、説明変数 (X) と目的変数 (y) から予測モデルを開発することである。目的変数の種類によって大別される。

- 回帰: 目的変数が（スカラー値として比較可能な）数値データ
- クラス分類: 目的変数が種類や分類を表すカテゴリカルデータ

機械学習では、より正解率の高い予測モデルを構築することが重要になる。高い正解率を実現するため、様々なアルゴリズム（手法）が考案されている。

- 重回帰
- 決定木、ランダムフォレスト
- サポートベクターマシン (SVM)
- ベイズ推定
- 多層パーセプトロン (MLP)

近年、人工知能ブームの立役者となっている深層学習は、多層パーセプトロンをより高度なネットワークモデルで構築したものである。よく使用されるネットワークには、CNN(畳み込みネットワーク)、RNN(リカレントネットワーク)があり、現在、Transformer など、盛んに新しいネットワークモデルが提案されている。

深層学習の利点として、従来の手法と比較すると、高い予測正解率が得られる反面、教師データを大量に集めることが重要になる。そのため、大量のデータが集まらない場合は、事前学習済み予測モデルを配布し、対象に合わせて転移学習する手法も積極的に用いられている。

3. 事例: 学習意欲の推定

データサイエンスや機械学習は、教育現場、特に ICT 利活用が進む教育現場において、膨大な教育活動のデータが蓄積されているため、様々な適用の可能性が考えられる。ここでは、我々が 2020 年度より試みてきた「オンライン・プログラミング演習において、学習意欲の低下している学生を発見する」事例を紹介する。

3.1 目的・動機

2020 年度は、COVID-19 の世界的な流行により、ほとんどの大学において、未経験の大規模なオンライン講義の提供となった。従来の対面方式の講義に比べ、学生の受講態度や反応から、学習意欲を把握することが難しくなった。

我々は、主に大学 2 年生に対して、Python による基礎的なプログラミング演習を行ってきた。これは、通常であれば、コンピュータ演習室で開講し、教員や TA が常時、質問などを受け付ける形式であった。演習室のやりとりから、つまづいている学生や意欲が低下している学生を把握することができた。しかし、オンライン・プログラミング演習になって、学生の状況を把握できず、一方向の演習になり、つまづいている学生を放置し、学習意欲を低下させることが懸念であった。

データサイエンスや機械学習の分析手法を用い、つまづいている学生を早期に発見し、適切な学習支援を行うことを目標として、学習意欲の低下によるドロップアウトを回避したいと考えた。

3.2 データ

2020 年度、突然のコロナ禍という状況下で、ほとんど事前に準備を行うことができず、オンライン演習の開講が決まった。利用可能であったデータは次の通りである。

- 学習ログ: 教育システムのアクセス記録
- 自由感想文

我々は、プログラミング演習の教材としてオンラインジャッジを用いている。2020 年度は、会津大学が提供する AOJ(Aizu Online Judge)、2021 年度は、AtCoder 社が提供する AtCoder を用いていた。どちらも登録されたユーザが提出したプログラム課題の実行結果が学習ログとして記録される。プログラム課題の正誤の判定は、プログラム実行時の入出力のみでチェックされ、ソースコードの内容は成績に評価されない。

図 1 は、オンラインジャッジから Web API を経由して回収した学習ログ（抜粋）である。ログは、課題 1 問ごとに記録され、正答 (AC) 以外に、誤答 (WA)、RA(実行時エラー)、TLE(時間超過)などが記録されるため、精密な学習ログとなる。ただし、正答例は Web 上で公開されているため、正答例を丸写しするなどの不正行為が可能であり、かつ、その不正行為を判定することは難しい。

自由記述の感想文は、毎週、演習時間のあとに学生に報告させた感想文である。我々は、大学が採用している LMS manaba を利用し、字数制限なしに書き込んでもらい、それらを集計した。manaba 以外にも、双方向のコミュニケーションツール Slack による質問を受け付けていたが、そちらの集計は行っていない。

なお、2021 年度も継続して同様な取り組みを続けている。2021 年は、準備をする期間は十分にあったが、結果と

id	name	date	problem	status
1				
2	Aaliyah	2020/5/8 19:50	ITP1_1_A	AC
3	Aaliyah	2020/5/8 20:11	ITP1_1_B	AC
4	Aaliyah	2020/5/8 20:14	335	AC
5	Aaliyah	2020/5/8 20:18	380	WA
6	Aaliyah	2020/5/8 20:25	380	AC
7	Aaliyah	2020/5/8 20:48	315	AC
8	Aaliyah	2020/5/8 20:50	ITP1_1_C	AC
9	Aaliyah	2020/5/8 20:58	ITP1_2_A	WA
10	Aaliyah	2020/5/8 20:59	ITP1_2_A	WA
11	Aaliyah	2020/5/8 21:02	ITP1_2_A	WA
12	Aaliyah	2020/5/8 21:05	ITP1_2_A	AC
13	Aaliyah	2020/5/8 21:07	ITP1_2_B	AC

図 1 学習ログ

id	date	comment
1	2020-05-07 11:52:41	プログラミングは正直苦手ですが、引き続き、なんとか問題を解けるように頑張っています。
2	2020-05-13 12:39:57	よるしくお願います。
3	2020-05-26 2:33:29	わからない問題が2問解決できたので、進歩できたと思います。
4	2020-06-02 17:50:23	借金の問題で一旦やる気を無くしましたが、ひたすら飛ばして先に進んでいます。いままでやってきたプログラミングの発想が役に立っていて、解けると楽しいです。
5	2020-06-09 13:03:36	わかんない問題が2問解決できたので、進歩できたと思います。
6	2020-06-16 21:03:03	難しくなってきた、正直少し嫌になってきましたが、自分で立てた最低限の目標はクリアしたいと思います。

図 2 自由感想文 (抜粋)

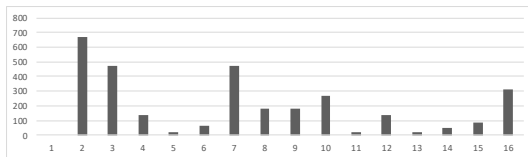


図 3 事例:コンスタントに学習したケース



図 4 事例:ドロップアウトしたケース

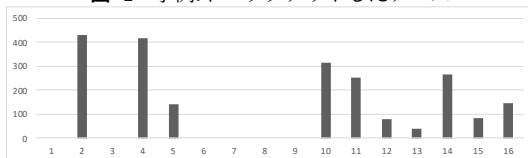


図 5 学習意欲を失わず復活したケース

して収集しているデータはほとんど変わらない。

3.3 学習ログ

まず、学習ログからオンライン演習の傾向を確認しておきたい。図3は、毎週欠かさず演習に取り組んだ学生の演習時間の推移を示している。演習室と異なり、オンライン演習では演習時間に対し強制力がないため、このようにバラツキが生じやすい。図4は、我々が早期に発見したいと考えているドロップアウトした学生のケースである。図5は、一見、ドロップアウトしそうに見えても、学習意欲は

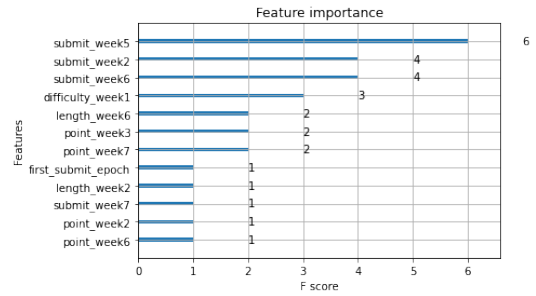


図 6 学習ログの影響の大きい特徴量

失われておらず最後まで演習を続けるケースである。

学習意欲は、学習ログから可視化すると、図3~図5のような時系列のパターンになる。我々は、「スポーツジムの退会者予測」の事例を参考に、決定木に基づく2クラス分類問題として、以下の通り、予測モデルの構築を行った。

- 時系列データ (日時秒単位) を一週間をビン (bin) として集計する
- 学生名と週単位でクロス集計し、提出回数、正答数とその累計を説明変数に用いる
- 最終的な課題提出量下位 25% をドロップアウトの候補者として、目的変数とする

我々は、単純な決定木アルゴリズムの代わりに、勾配ブーストを用いた決定木 (GBDT) XGBoost[1] を用いた。図6は、XGBoostにおける説明変数 (特徴量) の影響を調べたものである。第5週に提出した回数が最も影響を与えていることがわかる。

学習ログの解析で難しいのは、目的変数のラベル付けである。我々は、学習意欲の低下を3週以上の連続した欠席などから抽出したこともあったが、現在は、最終的な提出課題数に基づいてラベル付けしている。つまり、不正解であっても多くの課題を提出した学生は最後まで学習意欲を保ったとみなした。このようにすることで、高い正解率 (0.9 以上) も得られることが確認された。

3.4 自由感想文

自由感想文は、学習意欲を予測する上で、学習ログを補足するデータになる。

我々は、自然言語処理の代表的な手法である感情分析を参考に、自由感想文の特徴量をベクトル化して説明変数にして、スコア化された学習意欲を目的変数として、予測モデルの構築を紹介した。特徴量のベクトル化は、事前学習済み言語モデル BERT[2] を用いることで、限られたデータ数でも精度の高い予測モデルを得られるようにした。

自由感想文から読み取る学習意欲は、読み手によって大きく感じ方が異なる。例えば、「頑張ります」という感想から、まだ意欲を失っていないと感じるか、もう意欲を失っているとみなすのかは、一定するところではない。

当初、客観性の高いスコアを得るために、未提出回が3

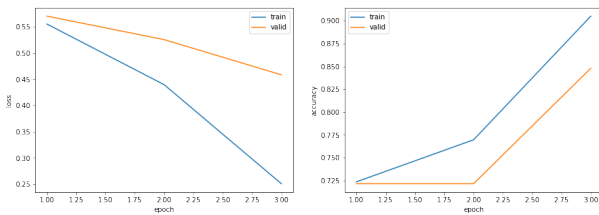


図 7 自由感想文の学習経過（損失と正解率）

回以上続くことを学習意欲の低下としてラベル付けし、学習モデル [3] を構築した。しかし、正解率は 0.72 とあまり高くならず、学習途中の損失も増減するなど、機械学習の適用としても望ましくない挙動を示した。

現在は、教員自身が感想文を読んで人手で学習意欲をスコア化し、予測モデルを構築する方向に変更した。その結果、図 7 のように、損失も正しく低下し、正解率は 0.85 と大きく改善がみられた。ただし、これはある教員の経験を学習したことになるが、本当に学習意欲が低下しているのか、その妥当性の検証には至っていない。

4. 知見と議論

我々は、オンライン・プログラミング演習における学習意欲の予測モデルを構築するため、学習ログと自由感想文の 2 つの収集データを解析している。本節では、少し視点を広げて、教育現場で教員がデータサイエンスや機械学習の活用を広げていくための知見を論じていきたい。

データ収集：データサイエンスや機械学習は、データ処理が中心的作業となるため、効率よくデータを収集することも重要になる。この点に関して、既存の教育システム (LMS) の対応状況はまちまちである。もし、LMS が、Web API によってデータを提供すると、分析タスクを自動化、リアルタイム化するのに役立つ。一方、LMS が Excel 形式のダウンロードだけサポートしていると、継続的なデータ分析を行うのは負担が大きい。今後 LMS の選択では、データ分析用の API を提供するかどうか、重要な項目になると思われる。

データ量：通常の一クラスで集められるデータ量は限られてくる。機械学習アルゴリズムによっては、十分な量に達しない可能性が多い。もし、予測モデルが期待通りの成果が得られなかったとき、データ量が足りないからか、それともデータ分析の仮説が間違っているのか、判断することが求められる。また、あわせてデータ量の少ないケースでも、事前学習モデルの活用、データの匿名化による複数データの統合など、予測モデルを構築することが求められる。

アルゴリズムと事例：現在、人工知能の民主化が進み、データサイエンスや機械学習に関する技術資料は、一般にも手に入りやすくなっている。しかし、データサイエンスや機械学習の多くの事例は、ビジネスや医療関係を題材

にしたものであり、教育分野の適用事例は意外に少ない。我々は、学習意欲を推測するモデルを作るため、スポーツジムの退会者予測や Twitter 等の解析で用いられる感情分析の手法を応用した。しかし、どちらも教育分野の前提と少し異なる部分があり、本当に適切な手法やアルゴリズムを選んでいるか議論になった。

教員が分析は可能か？本研究では、Python プログラミングに習熟した情報系学科の学部生や大学院生が中心になって、予測モデルの構築を行うことができた。このように、データサイエンスや人工知能の専門家でなくても、データ分析に取り組むことは可能といえる。しかし、一般の学校教員が同じ分析作業を行えるかと言えば、正直なところ、高い壁が残っているといえる。また、教員以外が分析をするときは、個人情報保護のための匿名化処理を行う必要があった。

今後、教育分野において、人工知能の活用が広がることは確実と思われる。ここで、技術力に優れた企業が人工知能を開発しても、その過程はブラックボックスとなる。教育現場では、人工知能は教員の評価の代行の面もあるため、ブラックボックスでは、納得感が少なくなると思われる。

最後に、本稿のまとめとして、教育 AI をブラックボックス化しないため、次のように教訓をまとめておきたい。

- 教育分野の事例を集めて公開すること
- 安全で使いやすいデータの匿名化を用意すること
- クラスや年度を超えたデータの大規模化を検討すること
- 事前学習済み予測モデルを用いて少量データでも精度を高めること
- 転移学習によるモデルのローカル化で納得感をあげる

5. まとめと展望

コロナ禍で突然にオンライン講義が始まったことにより、教育分野における ICT 活用が一段と加速した。ICT 活用が進むことで、より多くのデータが得られ、データサイエンスや機械学習の発展とともに、人工知能を活用した教育支援も期待される。一方、データをどのように教育支援に活用したかという具体的な事例は多くない。本稿は、学習ログや自由感想文を解析し、学習意欲の予測モデルを構築した経験を元に、今後、教育分野で人工知能を活用するときの知見を述べた。

なお、学習意欲の予測モデル自体は、経年のデータの集積や機械学習手法の検討が必要となる。今後、継続的に分析を進めて、さらなる報告していきたい。

参考文献

- [1] Tianqi Chen and Carlos Guestrin. Xgboost. *Proceedings of the 22nd ACM SIGKDD International Conference on*

- Knowledge Discovery and Data Mining*, Aug 2016.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [3] 田村 みゆ, 杉内 ゆり乃, 縫嶋 慧深, 田口 真里, 秋信 有花, and 倉光 君郎. オンライン演習下における自由記述感想文からの学習意欲の推論モデル. In **情報教育シンポジウム論文集**, volume 2020, pages 151–154, dec 2020.