

生成型 AI を用いた学習アドバイジングの提案及び評価

釣部 勇人^{*1}, 高野 泰臣^{*1}, 上野 春毅^{*2}, 小松川 浩^{*1}

^{*1} 公立千歳科学技術大学大学院 理工学研究科, ^{*2} 公立千歳科学技術大学 理工学部

A Proposal and Evaluation of Learning Advising using a Generative AI

Taketo Tsurube^{*1}, Yasuomi Takano^{*1}, Haruki Ueno^{*2}, Hiroshi Komatsugawa^{*1}

^{*1} Graduate School of Science and Engineering, Chitose Institute of Science and Technology

^{*2} Faculty of Science and Technology, Chitose Institute of Science and Technology

In this paper, we propose a method for automatically generating learning advice using a large language model, and evaluate that. In our previous study, advising message were created for each cluster based on learner classification using online learning history data. In the present study, we indicate that introducing a generative model allows us to create automatic advice using a generative AI that is combined with the learner characteristics within the cluster.

キーワード: 自然言語処理, 人工知能, アドバイス生成, 大規模言語モデル

1. はじめに

近年の学習支援教育では、生涯学ぶ習慣や主体的に考える力を持ち、予測困難な時代の中でどんな状況にも対応できる多様な人材の育成が求められている。こうした人材の育成には学習者が自ら学習を進め学習状況を振り返ることができる主体的かつ自律的な学習が重要となる。こうした中、高等教育では主体的に学ぶ姿勢の育成が求められている。学生主体の学びの中では、学習者自らが何をどう学ぶかを考えながら学習することが求められる。

しかしながら、学生主体の学習に慣れていない学習者は自ら学習を進めることが難しく、主体性を身に着けることが困難である。本来、こうした学習者に対して教員が学習状況や振り返り等から学習者の状況を察し、支援することが求められるが、学生の状況を毎週確認することは難しい。

一方、近年では大規模言語モデル(Large Language Model:LLM)の台頭により、生成型 AI への注目が高まっている。LLM は高水準の言語理解力及び表現力を有することで知られ、従来の小中規模言語モデルに比べ

てより柔軟な入出力が可能とされる。特に ChatGPT(GPT-3.5⁽¹⁾)や GPT-4⁽²⁾に代表される対話型の LLM は、自然言語によるタスクの指示が容易であることから様々な分野での活用が進められている。本研究においては、生成モデルを自動アドバイジングに取り入れることで、限られた出力から選択的にアドバイジングを行う分類モデルに比べ、より個々の学習者の特性に沿ったアドバイジングの実現が期待される。

2. 目的

本研究では、オンライン学習で蓄積可能な LMS(Learning Management System)の学習ログや振り返り等の情報を基に学習者の状況に合わせた自動アドバイジングを生成型 AI によって実現可能であるか検証する。対話型 LLM に学習者の情報をプロンプトとして与え、出力結果が学習者の主体性を養うためのアドバイスとして適切であるか確認を行う。学習ログや振り返り等の情報に基づくクラスタリングの手法を用いて学習者の傾向と比較して自動生成したアドバイスを評価する。

3. 自動学習アドバイジング

3.1 授業モデルの概要

自動アドバイジングの対象とする授業モデルを図 1 に示す。本授業モデルはオンライン型の反転授業モデルであり、学習者は予習内の全ての学習活動を非同期型で行う。これによって学習者が主体的に行動する活動が増加するため、自身の学習状況を把握しながら次週の目標と学習計画を立てられるように、各授業終了時に自己調整学習⁽³⁾を意識した振り返り及び目標設定を行う。グループワークは、Web 会議ツールを用いてオンライン上で実施する。また、授業への理解度が低い学習者や不安がある学生に対しては、知識の補完を目的とした講義形式の授業（ベーシッククラス）を行う。この授業モデルにおいてアドバイジングを各週の振り返り学習終了後に行い、次週の目標設定へと続けることで、自己調整学習における自己省察段階から予見段階への移行の際に、学習活動の妨げとならずにアドバイジングが行われる。

3.2 提案手法

学習者の学習履歴や振り返り情報を LLM へのプロンプトとして与えてアドバイジングの生成を行うことで、各学習者の特性を考慮した自動アドバイジングを行う。アドバイスの生成に用いる学習履歴及び振り返り情報を表 1 に示す。獲得方略及び回避方略の使用頻度の 2 軸による主体性の分類⁽⁴⁾を参考に、図 1 授業モデル内で収集可能な学習履歴のうち 6 項目を学習者の主体性に関する情報として用いた。各項目の値は数値や段階評価であり、すべて文字列として扱う。特に、段階評価

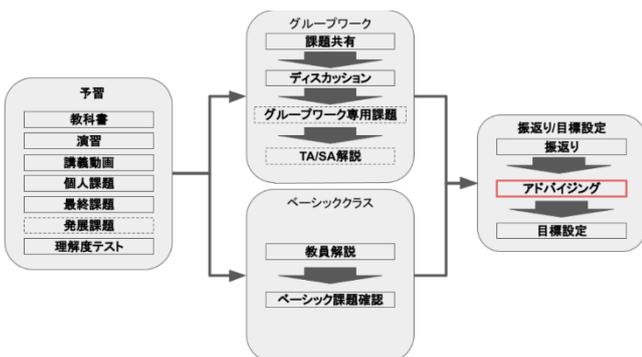


図 1 アドバイジングを踏まえた授業進行フロー

表 1 利用する学習履歴及び振り返り情報項目

	方略	項目名
学習履歴	獲得方略	教科書閲覧時間 演習取り組み回数 ワークシート取り組み状況 理解度テストのスコア
	回避方略	グループワークへの参加 グループワークへの貢献度
振り返り情報		1 週間前に立てた今週の目標 今週の目標の達成度 達成度の理由

は評価を示す数値ではなく、評価に紐づく文章を用いる。これらに加えて、学習者が自ら記述した 3 項目の振り返り情報を LLM へ入力する学習者の情報とした。

LLM へのプロンプトには、対話モデルの役割、学習者の情報(表 1)、出力フォーマットが含まれる。対話モデルの役割では「学習者を自己調整学習者にするために、適切な目標設定に関するアドバイスをする」よう指示しており、アドバイジングに用いる事前情報やアドバイジングの観点等の指定を行う。学習者の情報は、各学習者の項目データのみでなく、他学習者を含めた平均値や項目の補足を含み、箇条書き形式で情報が入力される。各学習者の情報以外の補足情報をプロンプトに含めることで、各項目を解釈するための基準を LLM に与えることができる。出力フォーマットは、生成されるアドバイスの出力形式を定めるものであり、「はじめに」「現在の状況」「今後の伸びしろ」「来週の目標設定に向けて」の 4 種のテーマについてアドバイスを生成するよう指示を行う。

4. 評価方法

本学の 2021 年度講義「C プログラミング」の 7 週目の授業データを対象として 90 名の学習者の学習履歴及び振り返り情報を入力し、LLM による自動アドバイジングの検証を行う。アドバイス文の出力には、OpenAI 社が提供する対話型の言語モデル gpt-3.5-turbo (temperature=0.7) を用い、出力結果に対して学習者の

分類結果との比較を行った。

学習者の分類では前述の学習履歴から得られる6項目を用いたK-meansクラスタリングにより4種のクラスタを作成する。作成されたクラスタはそれぞれ以下のような特性をもつ。

順調型 教科書や演習の取り組みが高くはないにもかかわらず、ワークシートの取り組み状況及び理解度テストのスコアが高水準であり、授業以外の教材も取り入れた主体的な学習が十分に行われている。

教科書やりこみ型 教科書閲覧時間が突出して高く、順調型に次いでワークシート及び理解度テストのスコアが良好である。

演習やりこみ型 演習取り組み回数が突出して高いが、グループワークへの参加及び貢献度が低くアクティブ・ラーニングに対して回避傾向にある。

伸び悩み型 ワークシートの取り組み状況及び理解度テストのスコアが良好でなく、教科書や演習の取り組みが十分でない。

各クラスタと特徴量の関係を表2に示す。表中の記号は各項目の達成度を示しており、「◎、○、△、×」の順に達成度が高い。自動アドバイジングでは、表2のような特性を持つ学習者に対して達成度が低い項目の改善を促すアドバイスの生成が求められる。また4種の分類において、特に主体的な学習者の特徴がみられる順調型を最も良好な状態であるとし、順調型への遷移を促すアドバイジングが期待される。

表2 各クラスタの学習者特性

到達段階	教科書	演習	ワークシート	理解度テスト	GW参加率	GW貢献度
順調型	△	△	◎	○	◎	◎
教科書やりこみ型	◎	△	○	○	○	○
演習やりこみ型	△	◎	△	◎	△	△
伸び悩み型	△or×	△or×	△or×	×	×	×

教科書閲覧時間：00:00:00/週

演習問題取り組み回数：12回/週

個人課題・最終課題の取組状況：全て解けた

理解度テストのスコア：5

グループワークへの参加：参加した

グループワークでの貢献度：あまり貢献できなかった

1週間前に立てた今週の目標：自分一人で課題の大きな部分のソースコードを書けるようになる
自分のソースコードを簡潔に分かりやすく説明できるように準備する。

今週の目標の達成度：達成できなかった

その達成度になった理由：内容や動く仕組みは理解することができたが、どのようにソースコード書けば良いのかが分からなかったから。

【はじめに】

素晴らしい1週間の学習履歴と内省ですね。良いところを続けていくと、さらに成長できますよ！

【現在の状況】

教科書閲覧時間はまだ少ないですが、演習問題に積極的に取り組んでいるのが良いと思います。グループワークにも参加しているので、チームワークを意識することも大切ですね。また、理解度テストでレベル5を取ったのは素晴らしいです！

【今後の伸びしろ】

教科書の閲覧時間をもう少し増やすことができると、より深く理解ができます。また、グループワークでの貢献度については、自分のアイデアを積極的に出すことで、もっと貢献できるようになるでしょう。

【来週の目標設定に向けて】

今週の目標が達成できなかった理由が分かっているのは良いことです。来週は、ソースコードを書く際にどのような手順が必要かを調べて、自分で書けるようにトライしてみてください。また、教科書の閲覧時間を増やし、理解度テストでレベル6以上を狙ってみましょう！

図2 自動アドバイジング成功例

5. 結果と考察

5.1 自動アドバイジング成功例

LLMによって生成されたアドバイスと学習者の情報間の矛盾やプロンプトの指示に従わないアドバイジング等の不適切な箇所を含まない出力を成功とし、多くみられた成功例を図2に示す。

図2の学習者は、教科書閲覧時間、グループワークでの貢献度等の達成度が低い。生成されたアドバイスは「教科書の閲覧時間をもう少し増やす」「自分のアイデアを積極的に出すことで、もっと貢献」等、達成度の低い項目の改善を促している。また、「来週は、ソースコードを書く際にどのような手順が必要かを調べて」という自由記述の振り返り情報への言及がみられる。他成功例においても、自由記述部分に対する言及を含む

アドバイスが出力されており、全体のうち「1週間前に立てた今週の目標」に言及したアドバイスは約42%、「達成度の理由」への言及は約52%であった。また、どちらかへの言及が含まれるアドバイスは約69%であり、学習者の自由記述による振返り情報を考慮したアドバイジングが十分に行われている。このことから分類では困難であった学習者の細かな差異に対応したアドバイジングの可能性が確認できた。

5.2 自動アドバイジング失敗例

各出力と特徴量を比較したとき、生成されたアドバイスのうち約36%に不適切な箇所が含まれていた。その中でも多くみられたアドバイジング失敗例を図3に示す。

教科書閲覧時間：00:00:00/週
 演習問題取り組み回数：169回/週
 個人課題・最終課題の取組状況：半分未満しか解けなかった
 理解度テストのスコア：7
 グループワークへの参加：参加していない
 グループワークでの貢献度：×
 1週間前に立てた今週の目標：今回よりもポイントについて詳しくなり、プログラミングの課題をこなせるようにする。
 今週の目標の達成度：達成できなかった

その達成度になった理由：ポイントの理解は、教科書を見たりベシッククラスで解説を聞いたことによって、最初の頃よりもできるようになった。しかし、いざ課題をこなすという時になると、エラーがでて中々進められないという状況になってしまっていた。

【はじめに】
 今週もプログラミング学習に取り組んでいたこと、素晴らしいですね！続けることが大切です。

【現在の状況】
 教科書の閲覧時間や演習問題への取り組み回数は、**平均以上**に行われていたことが伺えます。ただ、個人課題・最終課題の取組状況が半分未満だったこと、グループワークへの参加ができていなかったことが、課題ですね。

【今後の伸びしろ】
 理解度テストのスコアが7に達していることから、ポイントの理解力は上がっています。ただ、エラーが出てしまうという状況があったようです。今後は、誤った答えを見つけた時には、その理由を考え、自分で解決する力を身につけることが大切です。また、グループワークに参加することで、他者との協力やコミュニケーションスキルも伸ばすことができます。

【来週の目標設定に向けて】
 今週の内省を踏まえて、来週の目標を立ててみましょう。個人課題・最終課題に取り組むこと、グループワークに参加すること、そしてエラーが出た時に自分で解決することを目標にしましょう。自己内省はとても大切な学習方法です。今週の内省を活かし、自分自身を成長させていきましょう！

図3 自動アドバイジング失敗例

本検証における教科書閲覧時間の平均値は00:11:46であり、図3の学習者の教科書閲覧時間は平均以下である。しかしながら、生成されたアドバイスでは「教科書の閲覧時間や演習問題への取り組み回数は、平均以上」とされている。この例に限らず、最高評価であるにもかかわらずそれ以上の成績を目指すよう促すアドバイスが出力される場合もあり、平均値や最大値との比較における矛盾が多く、数値比較に課題が残る。また、他失敗例においてプロンプトに含まれない先週の内容を参照したアドバイジングが確認されており、事実と異なる出力(hallucination)についても課題が残る。

5.3 分類結果との比較

自動アドバイジングによって学習者がアドバイスを受けた学習履歴内の項目の分析結果を表3に示す。表3の各値は項目に言及したアドバイスのクラスごとの割合を示しており、最終行はすべての学習者を対象とした場合の値を示す。表2と表3を比較すると、演習やりこみ型の教科書閲覧時間、理解度テスト、グループワークへの参加といった達成度の低い項目に言及したアドバイスが多い傾向にあることが分かる。伸び悩み型は、表2においてすべての項目の達成度が低いが、表3では教科書閲覧時間とグループワークへの参加への言及が他項目に比べて多い。全体の言及率についてもこれら2つの項目への言及率が高く、教科書による基礎知識の定着やグループワークへの積極的な参加に重点を置いていることが分かる。しかしながら、学習量が十分でない伸び悩み型の学習者は順調型への直接の遷移は困難であり、グループワークへの誘導は適切でない。これに対し、段階的な改善を目指したアドバイスを出力させる場合には学習履歴の変遷の考慮が想定されるため、学習者の時系列情報が必要となる可能性がある。

表3 自動アドバイジングでの各項目の言及率

到達段階	教科書	演習	ワークシート	理解度テスト	GW参加率	GW貢献度	学習者
順調型	74.00%	42.00%	26.00%	30.00%	26.00%	20.00%	50
教科書やりこみ型	33.33%	33.33%	66.67%	66.67%	66.67%	0.00%	3
演習やりこみ型	57.14%	28.57%	71.43%	28.57%	71.43%	28.57%	7
伸び悩み型	76.67%	36.67%	50.00%	43.33%	96.67%	13.33%	30
	72.22%	38.89%	38.89%	35.56%	54.44%	17.78%	90

学習履歴及び振り返り情報のどちらに基づいた自動アドバイジングが行われているかについての分析結果を表4に示す。各値は学習履歴や振り返り情報等に言及したアドバイスの割合を示しており、最終行は表3と同様である。表4より、全体を通して振り返り情報よりも学習履歴への言及が多い傾向にある一方、伸び悩み型よりも順調型において振り返り情報に基づいたアドバイジングが多いことが分かる。これは伸び悩み型が学習履歴に多くの課題を抱えているのに対し、順調型は学習履歴項目での課題が少なく、振り返り情報に学習者の抱える課題が示される場合が多いことが要因であると考えられる。このことから、学習履歴と振り返り情報のどちらにおいても、各学習者の情報の中で特に達成度の低い項目に基づいて自動アドバイジングが行われていることが確認できた。

6. まとめと今後の課題

本研究の提案手法では、各学習者の達成度が低い項目に対して改善を促す出力が確認でき、生成型AIによる学習者のデータに基づいたアドバイスの自動生成の可能性を示すことができた。また、分類手法では困難であるクラスタ内の詳細な差異についても自由記述への言及から確認できた。

しかしながら、不適切な箇所を含むアドバイスも多くみられ、生成型AIによる安定した出力は得られていない。また、伸び悩み型のように段階的な改善を目指す場合には学習者の時系列情報の挿入する必要性が生じる可能性もあり、プロンプトへ与える情報の精査は今後の課題となる。さらに、プロンプトに対する出力は使用する言語モデルのモデルサイズや学習データセットに大きく依存するため、アドバイジングに用いるモデルについても、今後検討する必要がある。

表 4 自動アドバイジングでの学習履歴及び振り返り情報の言及率

到達段階	学習履歴 の考慮	目標/振り返り の考慮	混合 アドバイス	適切な アドバイス	混合&適切なア ドバイ
順調型	96.00%	80.00%	76.00%	68.00%	56.00%
教科書やりこみ型	100.00%	66.67%	66.67%	66.67%	66.67%
演習やりこみ型	100.00%	71.43%	71.43%	42.86%	42.86%
伸び悩み型	100.00%	50.00%	50.00%	60.00%	33.33%
	97.78%	68.89%	66.67%	63.33%	47.78%

- (1) OpenAI: “Training language models to follow instructions with human feedback”, arXiv:2203.02155v1 [cs.CL] (2022)
- (2) OpenAI: “GPT-4 Technical Report”, arXiv:2303.08774v3 [cs.CL] (2023)
- (3) B・J・ジーマン: “自己調整ハンドブック”, 北大路書房 (2014)
- (4) 白澤秀剛: “アクティブ・ラーニング導入による学生の主体性獲得プロセスのモデル化に関する研究”, 科学研究費助成事業 研究成果報告書, (2019)