

授業資料に基づく演習問題の生成手法

A Method for Generating Practice Exercises Based on Instructional Materials

青野 圭汰^{*1}, 上野 春毅^{*2}, 小松川 浩^{*1}

Keita AONO^{*1}, Haruki UENO^{*2}, Hiroshi KOMATSUGAWA^{*1}

^{*1} 公立千歳科学技術大学大学院 理工学研究科

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

^{*2} 公立千歳科学技術大学 理工学部

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

Email:m2250010@photon.chitose.ac.jp

あらまし：本研究では、生成系 AI (ChatGPT) を用いて、授業後の理解度確認を目的とした演習問題 (確認問題) を自動生成する手法を提案する。授業で用いるワークシートおよび e ラーニング教材に基づき、出題形式や内容の整合性を保ちながら問題文を生成し、6 項目のルーブリックにより出力の質を評価した。評価の結果、出力には一定の品質が見られる一方、単元の特定の用語に偏った問題文が多く、出題の多様性が課題であることが示唆された。

キーワード：生成系 AI, RAG, 問題生成

1. はじめに

近年、高等教育においては、学生の主体的な学びを促すことを目的とした反転授業や予習中心の授業設計が注目されている。特に、e ラーニング教材や演習問題を活用した事前学習は、授業内での応用的な活動を可能にする基盤として広く採用されている。

しかし、学生が予習内容をどの程度正しく理解しているかを教員が把握することは難しく、つまずきの箇所や理解の深さを詳細に捉えるには限界がある。

本研究が対象とする授業では、反転授業のアプローチを採用し、学生はまず、e ラーニング教材を用いた予習とプレテストを行った上で、演習活動やデータ分析課題などに取り組んでいる。具体的には、Python の基礎文法から機械学習・統計手法 (重回帰分析, ロジスティック回帰, クラスタリングなど) までを段階的に学び、最終課題でのデータ分析につなげる構成を取っている。このように、データサイエンスの基礎から応用に至る一連のスキルを段階的に修得させる構成となっており、演習を通じて学生が主体的に考えながら実践力を高めることを狙いとしている。

しかし、仮にプレテストを実施したとしても、結果だけでは学習の定着度やつまずきの内容を十分に把握することはできない。そこで本研究では、教材に含まれる演習問題をもとに、授業後の学習内容の定着を図る「確認問題」を自動生成する手法の構築を目的とする。

このような確認問題は、教員にとっては授業設計における負担軽減や指導の個別最適化に貢献し、学生にとっても自らの理解の振り返りと補強を促す有効な手段となりうる。本研究は、反転学習における「予習→プレテスト→授業→確認」の流れの中で、理解度を多面的に把握できる支援の一端を担うものである。

2. 研究の目的

本研究の目的は、プレテストを受験済みの学生に対し、内容理解を補完するための確認問題を、予習

ワークシートの内容に基づいて自動生成することである。特に、生成される確認問題は、プレテストと同様の形式を維持することを重視している。これにより、学習の継続性を確保しつつ、学習内容の定着度を高める教材としての活用を目指す。

3. 関連研究

近年、生成系 AI (例: ChatGPT) を教育支援に活用する研究が盛んになってきている。たとえば、既存の教材からプログラミング演習問題を自動生成するシステム⁽¹⁾が提案されている。しかし、これらの多くは自由形式の問題生成に留まっており、実際の授業設計において必要とされる「授業の設定」や「授業内容」への配慮が十分ではない。

4. 研究の手法

本研究では、本学の e ラーニング教材である Solomon の演習問題をベースに、授業ワークシートを ChatGPT に渡して確認問題を生成した。まず、授業ワークシートを用意し、これを、OpenAI 社が提供する Assistants API の RAG ツールである file search を通じて ChatGPT に与える。次に、生成される問題が、プレテストと同様の出題・解答形式になるようプロンプト内で演習問題の例示を行った。

問題文は、内容的妥当性 (ワークシートの内容と一致しているか) および形式的整合性 (プレテストと同じ書式) を両立するように設計されており、確認問題として授業内での使用を前提としている。

今回使用した授業回は Python を使った重回帰分析の第一回であり、ワークシートに対応する演習問題は問題に設定された難易度が 7 段階のうち 1 から 3 までの問題群である。今回は、各レベルで 1 問のみを例示として与える場合と、各レベルから 1 問ずつ計 3 問を例示として与える場合の全部で 4 種類のパターンで検証を行った。また、検証の際には出力のプレをみるため各パターンで 5 回ずつ出力している。

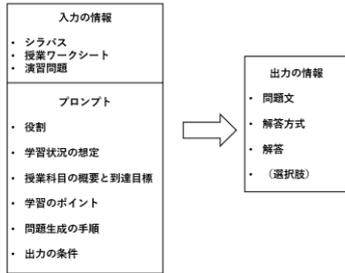


図 1：演習問題生成における入力情報と出力情報の構成

5. 検証結果

結果を以下の表 1 の項目に基づいて評価した。

表 1：評価項目

評価項目	3(真)	2(やや不十分)	1(不適切)
学習目標との整合性	学習目標と一致し、理解を的確に測れる	一部対応しているがずれがある	関連が薄い／ずれが大きい
問題の明確さ・文法	文法・構成ともに明快で読みやすい	表現があいまい／冗長	文法ミスが多く、意味不明
形式の正確さ	指定形式・選択肢構成が規定通り	形式は正しいが構成に乱れあり	指定と異なる形式／構造不備
内容の妥当性	前提や設定が論理的かつ明確	一部あいまいな点がある	論理の破綻、情報不足
難易度の適切さ	授業内容に対して適度な難易度	やや簡単 or 難しすぎる	極端に高い or 難解
解答の正確性	解答が内容に忠実に論理的	解答は正しいが説明が不足	誤答、または極端不明の解答

表 2：各問題における評価結果

問題	学習目標との整合性	問題の明確さ・文法	形式の正確さ	内容の妥当性	難易度の適切さ	解答の正確性
レベル1.1	3	3	3	1	3	1
レベル1.2	3	3	3	2	3	1
レベル1.3	3	2	1	2	3	1
レベル1.4	3	3	3	2	3	1
レベル1.5	3	3	3	2	3	1
レベル2.1	3	3	3	2	3	1
レベル2.2	3	3	3	2	3	1
レベル2.3	3	3	3	2	3	1
レベル2.4	3	3	1	2	3	1
レベル2.5	3	3	3	2	3	1
レベル3.1	3	3	1	2	3	1
レベル3.2	3	3	2	2	3	1
レベル3.3	3	3	3	2	3	1
レベル3.4	3	3	3	2	3	1
レベル3.5	3	3	2	2	3	1
レベル123	3	3	3	2	3	1
レベル123	3	3	3	2	3	1
レベル123	3	3	1	1	3	1
レベル123	3	3	3	2	3	1
レベル123	3	3	3	2	3	1

表 3：評価結果の項目別平均

条件	学習目標との整合性	問題の明確さ・文法	形式の正確さ	内容の妥当性	難易度の適切さ	解答の正確性
レベル1	3	2.6	2.6	1.5	3	1
レベル123	3	3	2.6	1.8	3	1
レベル2	3	3	2.6	2	3	1
レベル3	3	3	2	2	3	1

各パターンにおいて生成された問題に対して、6項目のルーブリックを用いて評価を実施した(表 1)。その結果、「学習目標との整合性」および「難易度の適切さ」はすべてのパターンで平均 3.0 と高い評価を得た(表 3)。一方で「解答の正確性」は全パターンで 1.0 と最低評価に留まり、最も大きな課題として明確に示された。

また、「形式の正確さ」や「内容の妥当性」についてはレベルによってばらつきが見られた。特にレベル 3 では「形式の正確さ」が平均 2.0 と低下しており、複雑な例示が出力の安定性に影響を与えた可能性がある。

さらに、レベル 1~3 を混在させた「レベル 123」パターンでは、形式の正確さは 2.6 と比較的高水準を維持していたものの、「内容の妥当性」は 1.8 と他パターンと比べてやや低下した。これは例示の多様性が出力の一貫性に一定の影響を与えることを示唆している。

6. 考察

生成されたすべての問題において、設問内容が「決

定係数を算出し、モデルの評価を行う」ものに偏っていた。これは、file search によって与えたワークシートの影響が考えられる。今回与えたワークシートでは、健康に関わるデータから重回帰による分析を試みる演習となっており、ワークシートの最後の演習で決定係数と重回帰係数の算出を行ってモデルの評価を行っている。ワークシート情報を参考にして問題作成するように指示をした結果、ワークシートの情報に生成される問題が強く影響を受けてしまったと考えられる。

このようなバリエーションの乏しさは、学習目標全体のカバー率の低下や、学習者に対する出題の単調化につながるおそれがある。今後は、出題例の提示内容の分散化や、学習目標に沿った多様な出題パターン(例：係数の解釈、交差項の意味、予測精度の比較など)を意識的に含めることが、より効果的な演習問題の自動生成につながると考えられる。

7. おわりに

本研究では、生成系 AI を活用して授業ワークシートに基づく確認問題の自動生成を試み、問題の形式的・内容的整合性を保ちながら出題可能であることを示した。評価の結果、全体としては高い整合性と適切な難易度が維持されていた一方、出題内容が決定係数に偏るなど、内容の多様性に課題が見られた。また、解答の生成に関しては、今回正しく生成できたものが 1 つもなく、解答の正確性についても課題が残された。

この要因として、生成系 AI は自然言語による問題文の生成には一定の強みを持つ一方で、数式処理や精密な数値推論を必要とする設問に対しては誤答や論理の飛躍が生じやすく、現時点では制御が難しいという特性が挙げられる。解答部分においては、一貫性や正確性を十分に担保することが困難であるという限界が明らかとなった。

今後は、出題例のバリエーション拡張やプロンプト設計のさらなる調整を通じて、より多様な学習項目に対応する問題生成を目指すとともに、出力された問題の質を評価するためのルーブリックの精緻化や、第三者による評価の導入も検討していきたい。また、解答の生成については、問題文の生成と工程を分離し、別フェーズでの検証・補完を行うことで、より正確で信頼性の高い教材構築につなげたいと考えている。

参考文献

- 前田 悠翔, 田中 英武, 井垣 宏, 福安 直樹: 多様な問題パターンおよび難易度を考慮した言語系生成 AI によるプログラミング演習問題自動生成手法の検討, 第 10 回実践的 IT 教育シンポジウム (rePiT2024) 論文集, pp. 88-99 (2024)