

テクノロジーと言語

吉岡峻佑 23B13565
東京工業大学 工学院

1. はじめに

現在、生成AIが注目を集めており、特に2022年に公開された“ChatGPT”がよく取り上げられている。各企業ではChatGPTの活用に向けて動き出している。

しかし、ここで問題となっているのがその生成された文章に対しての真偽についてだ。“ハルシネーション(幻覚)”と呼ばれ、嘘の答えをあたかも、もっともらしい答えのように発言することがある。

最近では子供の教育にてChatGPTを導入すべきか議論を呼んでいるが、果たしてChatGPTのは発言はどの程度信用していいのか疑問に感じた。そのため今回はChatGPTが教育界に参入できるに値するか考察していく。

2. 方法

日本学生の大学進学率は年々上昇し続けており、2023年では過去最高の57.7%に達した。そのため教育の目的、また学生の勉学に励む目的の一つとして大学進学があげられる。そのためChatGPTが大学進学にどの程度生かせるかを調べる。

方法として、誰でも無料で使用できるChatGPT-3.5を用いて、彼に大学の過去問題を解かせ、その正答率を調査する。過去問題は東京工業大学の2023年のデータを用いることにする。

3. 結果

科目	数学	英語	化学
正答率(%)	0(20)	58.3	7.6

表1: 各科目と正答率

各科目の正答率を上表1にまとめた。化学の一部の問題や物理に関しては図から読み解く問題が多く、調査の対象外とした。数学に関して(20)となっているが、ChatGPTでは複雑な数値計算によりChatGPT内で計算が難しいということで、プログラミング言語Pythonのコードが出力されたものがある。この出力されたコードをGoogle Colaboratoryにて実行し、返されたものが正解だったものの正答率となっている。

4. 考察

実験結果よりChatGPTが英語科目において約58%の正答率を示したことは注目すべき成果である。英語は言語理解が必要な科目であり、ChatGPTが過去問題に対して比較的高い正答率を達成できたことは、言語生成モデルの進歩を示唆している。しかし、他教科に比べて正答率が高いだけであり、ChatGPTを英語教育に活用するには程遠い数値である。

数学に関して、複雑な数値計算に対応できず、プログラミング言語を導入しても、20%の正答率というあまり優れていない結果となった。ChatGPTの仕組みとして、プロンプトされた内容に関して完全に理解したわけではなく、文脈からもっともらしい答えを推察して返答している。そのため、厳密な操作をしなくてはならない数学の問題を正確に解かせることは無理難題であり、ChatGPTの限界が現れたといえる。

また化学にてハルシネーションが他科目よりも多く確認された。化学の過去問題では化学知識を問う正誤問題が多く出題されているが、その多くが不正解であった。その間違った回答を選んだ根拠的の外れなことばかり書かれており、化学の根本的な理解には達していないことがわかる。

5. おわりに

今回の調査では生成AIであるChatGPTが教育分野での活用の可能性を探るため、大学入試の過去問題を対象に調査した。英語のように文章理解において一定の成果を上げた一方で、数学や化学のような理系科目では複雑な計算や抽象的な理論の理解に課題があった。

結論として、ChatGPTの利用には慎重な検討が必要である。教育分野では正確で信頼性の高い情報が求められる。ハルシネーションが多く確認される現在では、子供たちに誤った情報を伝える可能性が高い。将来的に活用するため、生成AIと教育界の連携を強め、継続的な研究が不可欠である。