

2023 年度 高等学院同窓会学術研究奨励金 研究成果報告書概要 (WEB 公開用)

高等学院長
高等学院同窓会理事長 殿

研究代表者氏名 [高橋 彰仁]

学年・組・番号 [1 年 B 組 26 番]

研究課題: クイズ生成アルゴリズムの開発

(英文) Development of quiz generation algorithm

研究概要:

(研究課題を選んだ動機、達成するための計画・目的・方法等について 200~400 字で記入してください)

既存のオンライン早押しクイズアプリにおける不満点の一つである、ユーザーが作成した特定のトピックの対戦が不足することを解決するアルゴリズムの開発を目指した。

人間と同質のクイズを生成する AI には、広範な知識、自然言語運用能力、創造性が必要である。そのため、これらの要素を備えた GPT-3.5/4 をアルゴリズムでの生成タスクに使用した。しかし、GPT-3.5/4 は、与えられたトピックについての具体的なクイズを生成できるほどの深く豊富な知識、またその知識を得るための情報収集能力を持っていない。

そこで私は、Web サイト上のテキストを収集する情報収集アルゴリズム、収集された情報を参照した GPT-4 がクイズを生成するクイズ生成アルゴリズム、収集した情報を 30 日間のみ保存して入力されたトピックに応じて保存した情報を使用する短期記憶アルゴリズム、不変的であるトピックのジャンルを半永久的に保存する長期記憶アルゴリズムを組み合わせたアルゴリズム「Barista」を考案した。

研究成果:

(研究の結果概要、結果に対するフィードバックや感想等について 200~400 字で記入してください)

Google Colaboratory と Google Custom Search API、GPT-4、GPT-3.5、Gemini Pro を使用して、クイズの多様さと質を向上させるため、情報収集、クイズ生成、短期記憶、長期記憶アルゴリズムの 4 つのアルゴリズムを動的に使い分けまた組み合わせる「Barista」と、このサービスをバックエンド用に改変した「Barista Public」を開発した。

その結果、「Barista」が生成するクイズは、人間による定性的な評価で 67%の精度を示し、「Barista Public」は「Barista」と同等のクイズ生成性能を維持しながら、処理高速化と軽量化、低コスト化、著作権回避を実現した。

また、本研究についての学術論文が「早稲田大学高等学院 論文・作品集 (2023 年度)」に掲載予定、ポスターは第 6 回中高生情報学コンテスト全国大会に出展される予定である。

「Barista」のソースコード: <https://colab.research.google.com/drive/1Mz8q181Gsp56kucyRKLtNtvtktiXMOeZ?usp=sharing>

「Barista Public」のソースコード: https://colab.research.google.com/drive/10rSQvRAaDd1raMUOX4ok69_0wTm_ka-z?usp=sharing

研究者: (以下の、代表者・分担者は学年・組・氏名を明記する)

研究代表者 1 年 B 組 26 番 高橋 彰仁

担当教諭 八百幸 大 教諭

(受給額: 30,000 円)

※研究課題、研究概要、研究成果、研究代表者名が WEB ページ上で公開されることに同意します

研究成果写真 :

(研究過程がわかる写真や、研究結果がわかる写真などを数点貼り付けてください)

大規模言語モデルを用いたクイズ生成アルゴリズムの作成

早稲田大学高等学院 1年 高橋彰仁

ソースコード : <https://colab.research.google.com/drive/1MZ8qi8IGsp56kucyRKLtNvtktXMOeZ?usp=sharing>



要旨

ユーザーが入力したトピックについての情報をGPT4/3.5とGoogle Search APIを用いて収集し、それらの情報からGPT-4にクイズを生成させるアルゴリズム「Barista」を作成した。その結果、人間による定性的な評価で83%の精度を示し、人間と同等の質のクイズを生成するアルゴリズムを作ることに成功した。

研究背景

2022年3月にOpenAIから人間と同程度の高度な言語運用能力を持つGPT-4が発表された。このGPT-4並びにGPT-3.5を活用して、情報を収集・保存し必要時にそれを参照してクイズを生成するアルゴリズムを作成できないだろうか。



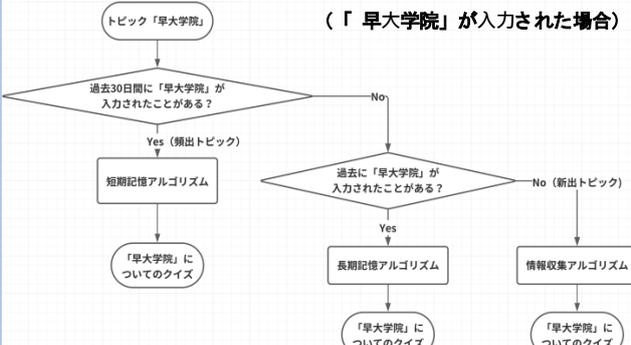
「Barista」の設計



「Barista」は情報収集/クイズ生成/短期記憶/長期記憶アルゴリズムの4つのアルゴリズムから成る。

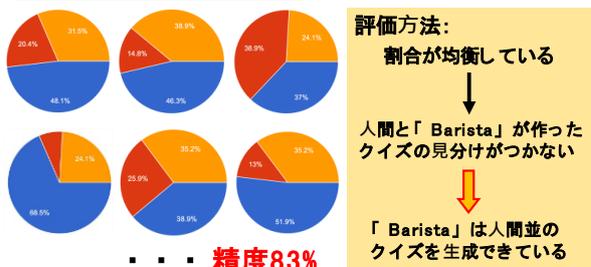
- 情報収集アルゴリズム:**
このアルゴリズムは入力されたトピックについての大量かつ深い情報を得ることを目的とし、人間の情報収集手順を模倣した以下の6つのステップを持つ。
 - 1-1. トピックの概要取得(by Search API)
 - 1-2. トピックに合うジャンルを選択 (by GPT-4)
 - 1-3. あらかじめ設定したジャンルに合ったキーワード (以下ジャンルキーワード) を検索クエリとして情報を収集 (by Search API)
 - 1-4. GPT-3.5が1-3で収集された情報を参照し、トピックに関連するキーワードを10個生成 (by GPT-3.5)
 - 1-5. 1-4で生成されたキーワードを検索クエリとして情報を収集 (by Search API)
 - 1-6. 1-3と1-5で収集した情報を結合する
- クイズ生成アルゴリズム:**
このアルゴリズムは収集データの保存、テキスト分割、クイズ生成を目的とし、以下の4つのステップを持つ。
 - 2-1. 1-6で結合した情報を shortterm_txtファイルに書き込んで保存 (後に短期記憶アルゴリズムで使用)
 - 2-2. 1-3で使用したジャンルキーワードをトピック名と共にLongterm.txtファイルに書き込んで保存 (後に長期記憶アルゴリズムで使用)
 - 2-3. 1-6で結合した情報をGPT-4の文字制限数以下になるように6900トークンに分割
 - 2-4. GPT-4が6900トークンのテキストグループを参照し、トピックについてのクイズを生成 (by GPT-4)
- 短期記憶アルゴリズム:**
このアルゴリズムは、過去30日間で入力されたことのあるトピックに対して機能する。まずShortterm_txtファイル内で入力されたトピックと同じトピックの情報がないか検索する。なければ長期記憶アルゴリズムが情報収集アルゴリズムが機能し、あれば短期記憶アルゴリズムが機能して、以下の3ステップのみでクイズを生成できる。
 - 3-1. Shortterm_txtファイル内の該当情報を抽出する
 - 3-2. 抽出した情報を6900トークンに分割する
 - 3-3. GPT-4が6900トークンのテキストグループを参照し、トピックについてのクイズを生成 (by GPT-4)また、Shortterm_txtファイル内の情報は古化を防ぐため、2-1で保存されてから30日後に自動削除される。
- 長期記憶アルゴリズム:**
このアルゴリズムは、過去に入力されたことのあるトピックに対して機能する。まず、Longterm_txtファイル内で入力されたトピックと同じトピックのジャンルキーワードがないか検索する。なければ情報収集アルゴリズムが機能し、あればこのアルゴリズムが機能して、1-1と1-2の処理をスキップできる。

「Barista」の構成



「Barista」はトピックの頻出度合いに応じて、短期記憶/長期記憶/情報収集アルゴリズムを使い分け、常に最小限のリソースでクイズを生成しているため、ユーザーが「Barista」を使用すればする程、記憶が増えて費用対効果が高くなるように構成されている。また、短期記憶内の情報は保存されてから30日後に自動削除され、削除されたトピックは入力時に長期記憶アルゴリズムに割り当てられて情報が再収集されるため、常に直近1ヶ月以内の最新情報に関するクイズを生成できるようにも構成されている。

生成されたクイズの評価



学校の生徒54人に、記載されたクイズ(四肢選択問題)が、人間が作成したクイズ、「～のクイズを作成して」と指示されたGPT-4が生成したクイズ、「Barista」が生成したクイズのどれなのかを選択してもらったアンケートを実施した。私は、人間と「Barista」を選んだ人の割合が均衡していれば「Barista」が生成したクイズは人間が作成したクイズと同等の質であると判断できると考えた。結果は6問中5問の割合が均衡していた。よって、「Barista」は83%の精度で人間と同等の質のクイズを生成できると言える。私個人も様々なトピックでクイズを生成してみて人間並のクイズを生成できていると感じた。
※「Barista」は四肢選択問題/記述問題/穴埋め問題/並び替え問題を生成できる。生成クイズ数は収集したテキスト量に応じて多少変化するが、約20問である。

結論・展望

83%の精度で人間が作成したものと同様のクイズを生成できていることから、「Barista」は人間と同等の正確性・具体性・自然さを持つクイズを生成できていると言える。しかし、「Barista」には処理時間が長い、高いコストがかかるという2つの問題がある。これらに対し、各タスク(ジャンル選択、キーワード生成、クイズ生成)に特化するよう、タスクに合った良質なデータセットでファインチューニングしたGPT-4/3.5よりもパラメータ数の少ない軽量かつ無料で使用可能なLLMをアルゴリズムに使用することで対処できると考えている。これができるれば、ファインチューニングによるLLMの推論時間の短縮、LLMの使用料がかからないことによるコストの大幅削減が実現可能だろう。今後は、「Barista」の精度向上/処理速度向上/低コスト化を進めていき、最終的には無料で手軽に好きなクイズを生成して楽しめるアルゴリズムにしたい。

参考文献

- 「Barista」ソースコード (Google Colaboratory)
<https://colab.research.google.com/drive/1MZ8qi8IGsp56kucyRKLtNvtktXMOeZ?usp=sharing>