



SIGNATE AI王

問題作成部門 予選3位

ELAIZO : kosirowada kkishikawaee

発表者

名前:和田幸志郎

所属:明治大学理工学部情報科学科

数理最適化研究室 – 主に機械学習などの最適化論

学年:B3

EpicAI – AI受託の会社の創業メンバー

企業向けのAI事業開発や生成AI導入支援など

mail:kosiro131219@gmail.com

手法: LLM + RAG を利用した問題生成

- ・LLM

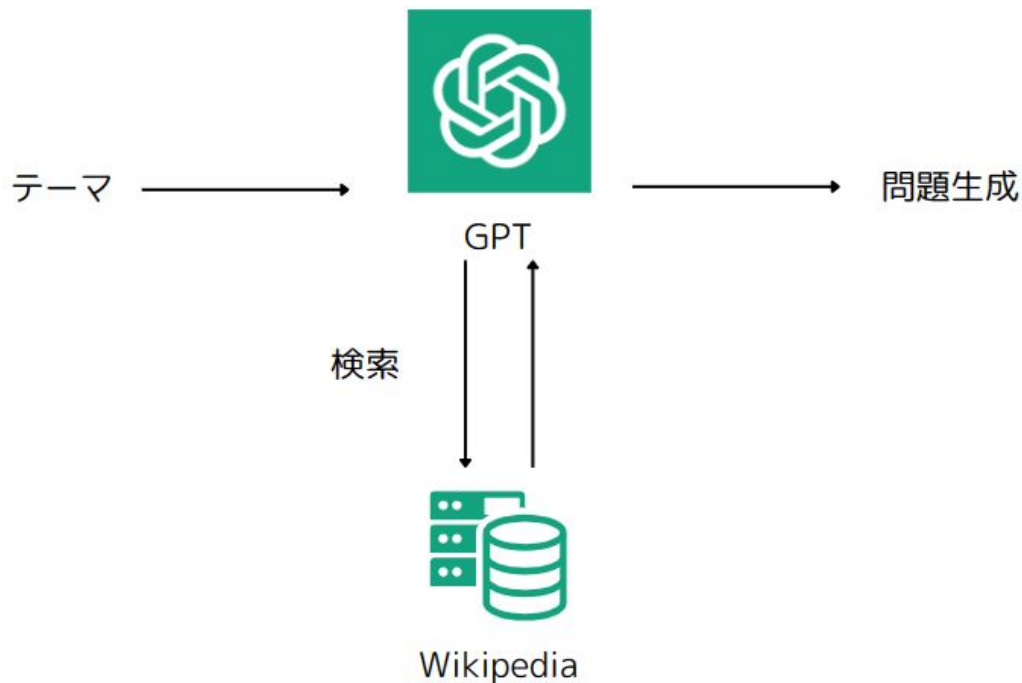
GPT-4を利用

- ・RAG

Wikipedia データの検索

- ・プロンプト

Slef-Refineを利用



RAG (Retrieval Augumented Generation)

RAG : 検索拡張生成

1. Wikipedia データベースからテーマを検索
2. 検索結果の文章をLLMに投げるプロンプトに入れる
3. LLMによる文章生成



RAG の例 (テーマ : サッカー)

プロンプトに検索結果を挿入

サッカーに関する問題を生成して下さい。

サッカーに関する情報は以下の通りです。サッカー（英: soccer）は、サッカーボールを用いて1チームが11人の計2チームの間で行われるスポーツ競技。足でボールを蹴る球技の総称であるフットボール（英: football)の一つであり、アソシエーション・フットボール（英: association football、協会式フットボールの意）またはアソシエーション式フットボール（略称：ア式蹴球 の略称である。他のフットボールコードと比較して、手の使用が極端に制限されるという特徴がある。... 続く

予選 (~11/30)

予選時は、GPT (text-davinci-003) を用いて、Wikipediaのデータセットから検索する RAG のシステムで提出。



「問題文と想定解が正しく対応しているか」という評価項目において、他のチームより低い点。(200点満点中 他チーム: 178, 147, 128, 130, 131, 27 自チーム: 97)



決勝では、「問題文と想定解が正しく対応しているか」を満たすような問題文を生成する必要がある。

予選時のプロンプト

f"以下の情報に基づいて、テーマ:{theme}に関する早押しクイズの問題文とその答えを作成してください。"

"問題は単語で答えることの可能な問題にしてください。"

"ただし、問題文は、「問題：」で始め、語尾は絶対に「〇〇でしょう？」という形にしてください。"

"また、答えは「答え：」で始めてください。...\n\n{retrieved_content}\n\n問題："

決勝

決勝での変更点

1. モデル (engine) の変更 : text-davinci-003 → GPT - 4
2. プロンプトの改善 : 評価指標をプロンプトに入れる
3. Self-Refine : LLMを使って初期出力し、その出力に対するフィードバックをし、それを使って繰り返し改良していく手法

プロンプトの改善

評価指標をプロンプトに入れてあげる

```
prompt = (  
    f"以下の情報に基づいて、テーマ:{theme}に関する早押しクイズの問題文とその  
    答えを作成してください。\\n"  
    "答えは必ず単語で答えられる問題を生成してください。\\n"  
    "ただし、答えは「答え：」で始めてください。...\\n\\n"  
    "問題作成のポイント：\\n"  
    "- 問題文から答えが一意に絞り込めるか。\\n"  
    "- 問題文と想定解が正しく対応しているか。\\n"  
    "- 問題文と想定解が正しい内容と言えるか。\\n"  
    "問題文は80字以内にしてください。\\n"  
    "もう一度言いますが、答えは「答え：」で始めてください。...\\n\\n"  
    "情報：{retrieved_content}\\n\\n"  
    "問題：" )
```

Self-Refine (プロンプト設計法)

1. LLM を用いた初期出力の生成
2. LLM による生成した出力の評価
3. LLM が評価を基に再度出力を生成



評価に関して

クイズらしい問題を生成するために、問題文中の語句に修飾をつけるように評価

例: (乃木坂46)

乃木坂46 のデビューシングルのタイトルは何でしょう？



日本の人気アイドルグループ、乃木坂46 の 2012 年にリリースされたデビュー
シングルのタイトルは何でしょう？

評価用のプロンプト

```
prompt = (  
    f"以下のクイズの問題と答えをレビューしてください。\\n\\n"  
    f"問題：{quiz_question}\\n"  
    f"答え：{quiz_answer}\\n\\n"  
    "このクイズの質はどうですか？ 問題や答えに改善が必要な点がありますか？\\n"  
    "レビューのポイント：\\n"  
    "- 問題の語句に修飾語を付けるとしたらどこに何を付けるか。\\n"  
    "- 問題文から答えが一意に絞り込めるか。\\n"  
    "- 問題文と想定解が正しく対応しているか。\\n"  
    "- 問題文に雑学的なトピックや興味深い情報などが適切に盛り込まれているか。具  
        体的にどんなトピックを入れるかを考えてください。\\n"  
    "ただし、問題文は複数になってはいけません。"  
    "また、解答も単語にならなければなりません。"  
    "補足情報は名詞を修飾する形で、必ず問題文に入れてください。"  
)
```

最終出力用のプロンプト

```
prompt = (  
    f"以下のレビューを参考にして最終的な問題と答えを作成してください。"  
    "ただし、答えは「答え：」で始めてください。...\n\n"  
    "また、問題文は複数の文にならないようにしてください。"  
    "日本語として違和感のない文章にしてください。"  
    f"テーマ：{theme}\n"  
    f"問題：{quiz_question}\n"  
    f"答え：{quiz_answer}\n\n"  
    f"レビュー：{review_result}"  
)
```

決勝戦を終えて

- ・20 問作成するのにかった時間: 1031 秒
- ・新しいトピックに対応できていない

例:「新しい学校のリーダーズ」、「蛙化現象」

- ・「シャネル」に関する問題のみ、生成に失敗していたため、再度生成
- ・問題を前から絞れるようにする

discussion (その他の手法の案)

LLM に問題を生成させる段階で、複数の問題を生成させる



Refine で、複数の問題から 1 問に絞る



その 1 問今回の Refine プロンプトを用いて評価



再度問題生成

参考文献・参考資料

各種プロンプト手法、Self-Refine

- [LLMのプロンプト技術まとめ #ChatGPT - Qiita](#)

ソースコード

- [kosirowada/AI_Problem_Generate_Competition: SIGNATEで開催された第4回AI王のリポジトリです。\(github.com\)](#)

コンテストページ

- [第4回 AI王\(問題作成部門\) | SIGNATE - Data Science Competition](#)