

Few-shot プロンプトを用いた近代小説の現代語訳の試み

立命館大学 韓毅弘, 青木志真, 安尾萌, Junjie Shan, 西原陽子

1. 研究の背景と目的

- 近代に書かれた文章には名作が多いが、読みにくいことが多い
- 事前に現代語訳したものを読むことで、理解の助けになる
- 近代に書かれた文章の読解を支援することを目的とする

2. 翻訳手法

- Few-shotプロンプトを用いてLLMで翻訳を行う
- 難しい英文の平易化を行った[Yutao, 2023]のプロンプトを参考にして、Few-shotのプロンプトを作成する

本研究のプロンプト

私の複雑な文章を簡単な文章に書き換えてほしい。意味は同じにして、もっとシンプルにしてください。

複雑：いくばくもなく官を退いた後は、故山、號略に帰臥し、
簡単：やがて役人を辞めた後は、かつて住んでいた號略（かくりやく、地名）へと戻り、
 ……

お手本となる回答例

複雑：袁俊はじめ一行は、意をのんで、叢中の声の語る不思議に聞入っていた。声は続けて言う。
簡単：{出力}

質問

Few-shotプロンプトの例

私の複雑な文章を簡単な文章に書き換えてほしい。意味は同じにして、もっとシンプルにしてください。

複雑：それは、思った通り、原稿用紙を綴(と)じたものであった。

簡単：それは思った通り原稿用紙をとじたものであった。

複雑：これをどうか、はっきりと、お覚えなすって下さいませ。

簡単：これをどうかはっきりと覚えておいてください。

複雑：むかし、いつの頃(ころ)でありましたが、竹取(たけと)りの翁(おきな)といふ人(ひと)がありました。

簡単：

出力例

🌀 簡単：昔、竹取の翁という人がいました。

翻訳の実行例

3. 検証実験

目的：

- 与えるペア数が翻訳精度に与える影響
- 与える原文と翻訳文のペアの文体の違いが翻訳精度に与える影響

表2 実験で用いた小説

作品名	作者	初版発行年	文数	文字数
山月記	中島敦	1942	178	6970
高瀬舟	森鴎外	1916	211	9359
人間椅子	江戸川乱歩	1925	287	16046
藪の中	芥川龍之介	1922	277	9712
竹取物語	和田萬吉	1928	162	12903

- 1 漢字多い
- 2 通常
- 3 仮名多い

実験内容：

- Few-shotプロンプトで与える原文と翻訳文のペア数を2, 3, 4個と変化させ、翻訳精度を算出した。
- 5つの小説の中から1つを選び、Few-shotのペアを選択し、5つの小説の翻訳をし精度を算出した。これを5回繰り返した。

評価方法：

翻訳結果を「翻訳された」「ほぼ翻訳された」「翻訳されなかった」の3段階で人間が判断し、それぞれの割合を算出

実験結果：

- ペア数を増やすほど1の割合が向上することが確認された。
- 文体が異なるペア間では1の割合が低くなる傾向が確認された

表3 ペア数ごとの翻訳精度

ペア数	1. 翻訳された	2. ほぼ翻訳された	3. 翻訳されなかった
2	0.624	0.207	0.169
3	0.630	0.203	0.168
4	0.647	0.193	0.159
平均	0.634	0.201	0.165

表4 与えたペアの違いごとの翻訳精度。翻訳されたものの割合を示す。左が与えられた原文、右が翻訳された文を表す

原文\翻訳	山月記	高瀬舟	人間椅子	藪の中	竹取物語	平均
山月記	0.614	0.586	0.686	0.581	0.580	0.609
高瀬舟	0.578	0.648	0.749	0.648	0.576	0.640
人間椅子	0.571	0.640	0.744	0.680	0.591	0.645
藪の中	0.569	0.596	0.707	0.639	0.595	0.621
竹取物語	0.559	0.678	0.718	0.668	0.640	0.653
平均	0.578	0.629	0.721	0.643	0.596	0.634

4. 今後の課題

- 文体の違いが翻訳精度に与える影響の考察
- 現代語以外の親しみやすい言葉への翻訳