
RAG [1] を用いたドメイン特化型 LLM [2] の構築および応用

1 はじめに

近年, 人工知能技術は急速な発展を遂げている.
その中で大規模言語モデル (Large Language Models:
LLM) が注目を集めている. 特に,

2 要素技術

2.1 RAG

2.2 データ

3 提案手法

4 実験

提案手法の有効性を確認するために,

4.1 実験データ

4.2 実験 1

4.3 実験 2

5 実験結果

5.1 実験 1 の結果

5.2 実験 2 の結果

5.3 考察

6 まとめと今後の課題

本研究では

参考文献

- [1] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks. 2021.
- [2] Masato Hirakawa, Shintaro Horie, Tomoaki Nakamura, Daisuke Oba, Sam Passaglia, and Akira Sasaki. elyza/llama-3-elyza-jp-8b. 2024.