

報告書

1 今週の進捗

- Text to Knowledge Graph に関する論文調査

2 論文調査

Knowledge Graph の作成により, 情報を体系的に整理して機械が理解可能な形に変換することができる. 具体的には以下のようなことが可能になる.

情報の整理 文章中に含まれる情報をグラフとして可視化

自然言語処理タスクへの応用 機械翻訳や質問応答システムにおいて情報を効率的に処理

機械理解 機械が文章を理解するための基盤

データ統合 複数の文章やソースから得られる情報を結合し, 異なるデータソース間の関係を表現

以下に調査した論文における Text to Knowledge Graph の目的をまとめる.

2.1 Knowledge Graph Generation From Text [1]

Knowledge Graph は推論, 意思決定, 質問応答など, さまざまなダウンストリームアプリケーションで活用できる. そのため, Knowledge Graph の自動構築によって豊富なテキストコーパスに存在する情報を整理, 構造化, 圧縮された形式で表現したい.

2.2 ReGen: Reinforcement Learning for Text and Knowledge Base Generation using Pretrained Language Models [2]

Knowledge Graph 表現は複雑な関係をもつノードエンティティ間の現実世界の情報を捉える強力なツールである. 自由形式のテキストから Knowledge Base を自動生成し, その逆の操作である Knowledge Base から適切なテキストを生成することは困難な研究テーマである.

2.3 From Text to a Knowledge Graph: The Information Extraction Pipeline

<https://neo4j.com/blog/text-to-knowledge-graph-information-extraction-pipeline/#>

Knowledge Graph を作成する際に直面する課題の 1 つは Knowledge Graph に組み込みたい情報の大部分がニュース記事, 電子メール, 科学雑誌のエントリーなどの非構造化テキストデータに存在することである.

3 今後の課題

- 関連論文の調査
- Knowledge Graph from Text の再現実験

参考文献

- [1] Igor Melnyk, Pierre Dognin, and Payel Das. Knowledge graph generation from text. 2022.
- [2] Pierre L. Dognin, Inkit Padhi, Igor Melnyk, and Payel Das. Regen: Reinforcement learning for text and knowledge base generation using pretrained language models. 2021.