

報告書

1 コードの使用許可のメールに関して

”rnayak@qmed.ca” に送信したが返事が来ない. メールアドレスが間違っている可能性があるが, 調べても他にメールアドレスが見つからないためなぜ返ってこないかわからない.

2 Text to KG モデル “Grapher”

テキストから Knowledge Graph を抽出するモデルとして Igor Melnyk らが Grapher [1] を提案している. 以下に Grapher の学習モデルおよびテストモデルの概略図を示す.

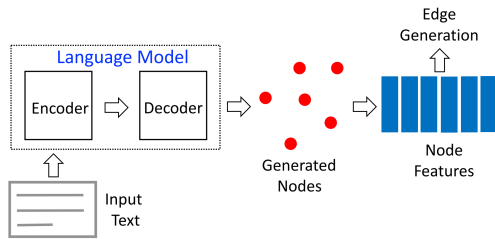


図 1: 学習モデル

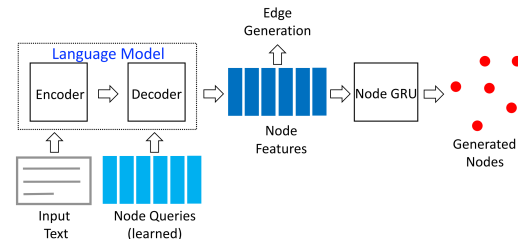


図 2: テストモデル

再現実験としてとりあえず動作して現在結果を待っている. また, コードおよびモデルの詳しい理解はできていない. 結果が出るのを待つ間に理解し, どのように改良するかを考えたい.

3 Text to KG の評価手法の調査

- WebNLG データセット
- TEKGEN データセット

いずれも文とその文から取れるトリプルの組. F1 値, Precision, Recall を用いて評価.

ホームズデータセットも文とトリプルの組のため, 同じ評価手法が可能である. しかし, 研究の目的が “物語の内容理解のためのナレッジグラフ生成” のため, データセット内のトリプルとの誤差をみると目的との相違が生じる. 他の評価手法を調査, 検討する.

3.1 WebNLG データセット

エントリ Id: 1

- 元のトリプル: Aarhus Airport — cityServed — “Aarhus, Denmark”@en
- 修正後のトリプル: Aarhus Airport — cityServed — “Aarhus, Denmark”
- 生成された文 1: “The Aarhus is the airport of Aarhus, Denmark.”
- 生成された文 2: “Aarhus Airport serves the city of Aarhus, Denmark.”
- コメント: “good”

4 今後

- Text to KG の評価手法の調査
- ホームズデータセットへの適用

参考文献

- [1] Igor Melnyk, Pierre Dognin, and Payel Das. Knowledge graph generation from text. 2022.