

報告書

1 今週の進捗

- 訓練データにおいて正解データのみを入力としたテスト結果
- 積み上げ棒グラフは未完成
- MLM を用いた実験コードは未完成

2 KG-BERT [1]

2.1 モデルの説明

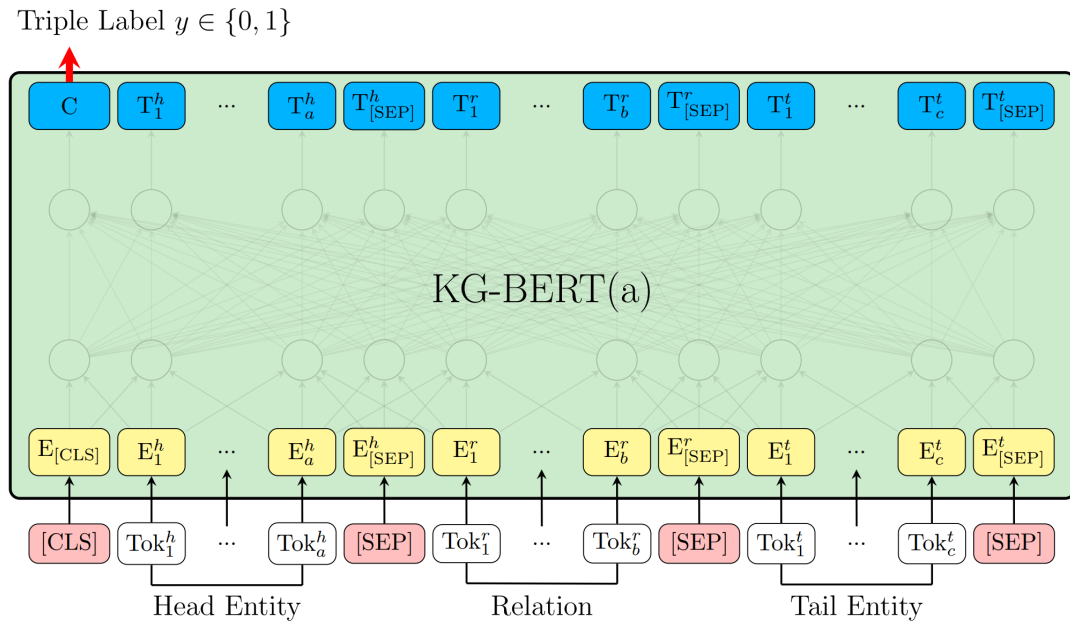


図 1: KG-BERT model [1]

表 1: データセット

Dataset	Entity	Relation	Train	Validation	Test
WN18RR	40,943	11	86,835	3,034	3,134

2.1.1 正解データだけでテストした結果

表 2 に訓練データとテストデータにおいて正解トリプルだけをテスト時の入力として実験した結果を示す。

表 2: 正解データだけの実験結果

	テスト数	正解数	Accuracy
訓練データ	86,835	86,794	0.9995
テストデータ	3,134	3,000	0.9572

2.2 エンティティ数 (40,943) 分類

2.2.1 存在するトリプルの数の分だけ上位にあるトリプルの割合

図 2 に head, relation が同じで tail が異なるトリプルの個数別における存在するトリプルを上位に予測した割合を表した結果を示す。横軸を head, relation が同じで tail が異なるトリプルの個数, 縦軸をトリプルの種類数としており, 青色の部分は正解したトリプルの種類数, 橙色の部分は不正解となったトリプルの種類数である。

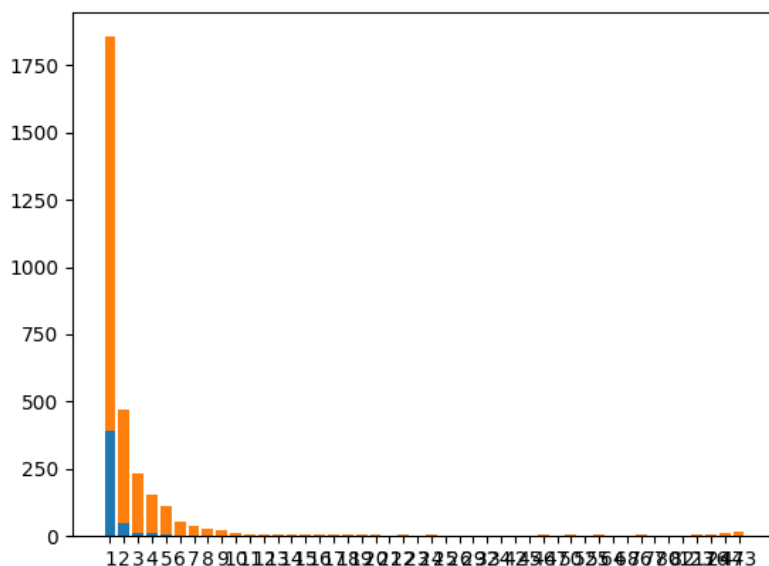


図 2: 全体の結果

図 3, 4, 5 に図 2 を 1 から 5, 6 から 10, 11 から最後までに分割した結果を示す。

2.3 MLM を用いた実験

現在, プログラムを書いている途中で, まだ完成していない。

3 今後したいこと

- KG-BERT に MLM を適用して実装
- エンティティ数 (40,943) 分類の改良
- KG-BERT の改良

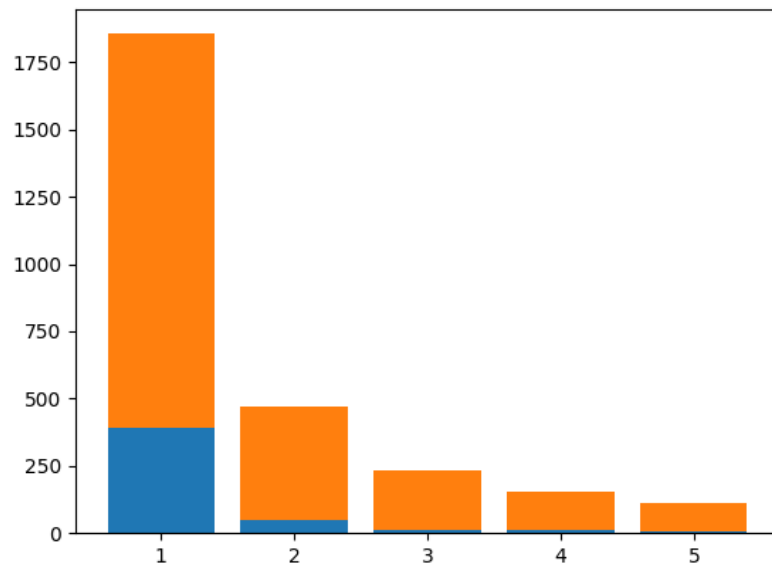


図 3: 1-5 の結果

参考文献

- [1] Liang Yao, Chengsheng Mao, and Yuan Luo. KG-BERT: BERT for knowledge graph completion. *CoRR*, Vol. abs/1909.03193, , 2019.

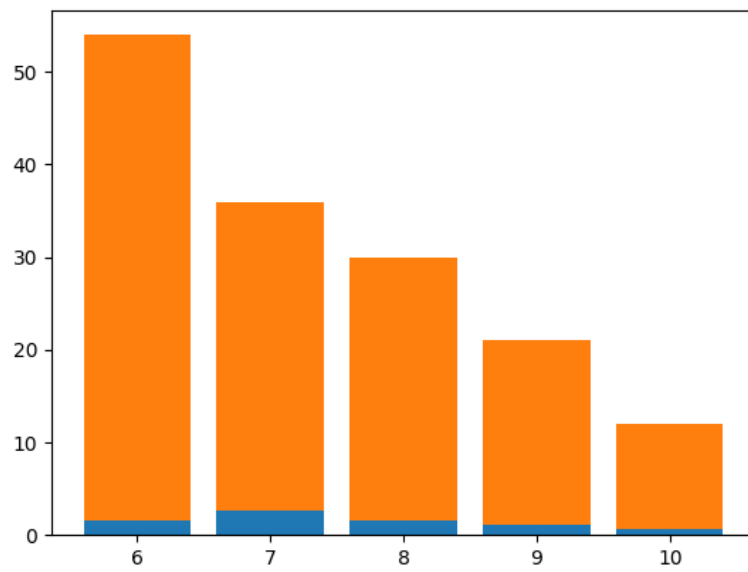


図 4: 6-10 の結果

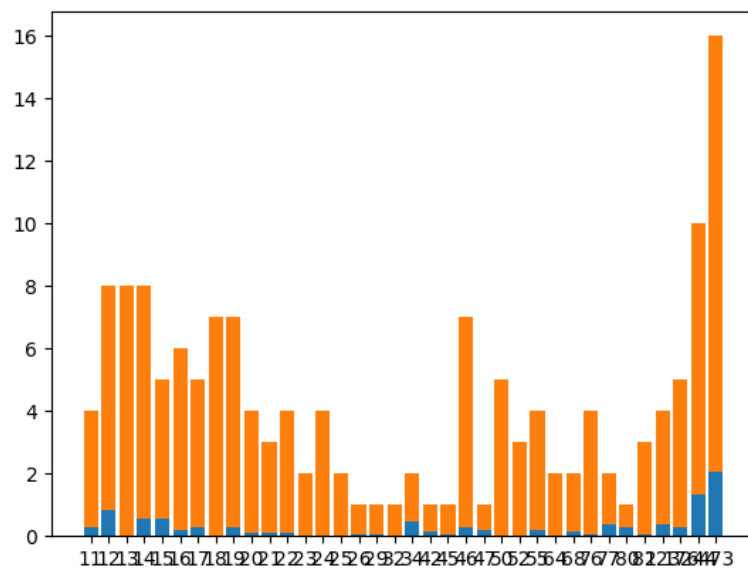


図 5: 11- 最後までの結果