

報告書

1 今週の進捗

- Relphormer の再現実験のやり直し
- KG-BERT の再現実験
- KG-BERT のコードの理解
- Relphormer の論文も KG-BERT の論文も arXiv にしかなかったです。

2 Relphormer [1]

実験をやり直したところ, 文献値には届かないものの前回の結果と比べると精度が上がりました。

前回の実験では訓練がうまくできていなかったみたいで, 訓練からやり直すと精度が上がりました。epoch 数を増やすと文献値に近づく可能性はあると思います。

表 1: 再現実験のパラメータ (WN18RR)

パラメータ		値 (変更前)	値 (変更後)
pre-train	学習率	1e-4	1e-4
	epoch	30	30
	num workers	8	8
	batch size	16	16
	eval batch size	16	16
	max seq length	32	32
test	学習率	3e-5	3e-5
	epoch	30	30
	num workers	4	4
	batch size	8	8
	eval batch size	8	8
	max seq length	16	32

表 2: 再現実験の結果

	WN18RR			UMLS	
	H@1	H@10	MRR	H@10	MR
変更前	0.0223	0.0616	0.0357	0.988	1.77
	0.0180	0.0517	0.0296		
変更後	0.114	0.315	0.180		
文献値	0.448	0.591	0.495	0.992	1.54

3 KG-BERT [2]

3.1 再現実験

今まわしています. 結果は明日には出ます.

表 3: パラメータ (WN18RR)

パラメータ	値 (default)
学習率	5e-5
epoch	5
dropout rate	0.1
batch size	8 (32)
eval batch size	5000
max seq length	32 (50)

3.2 KG-BERT のモデル

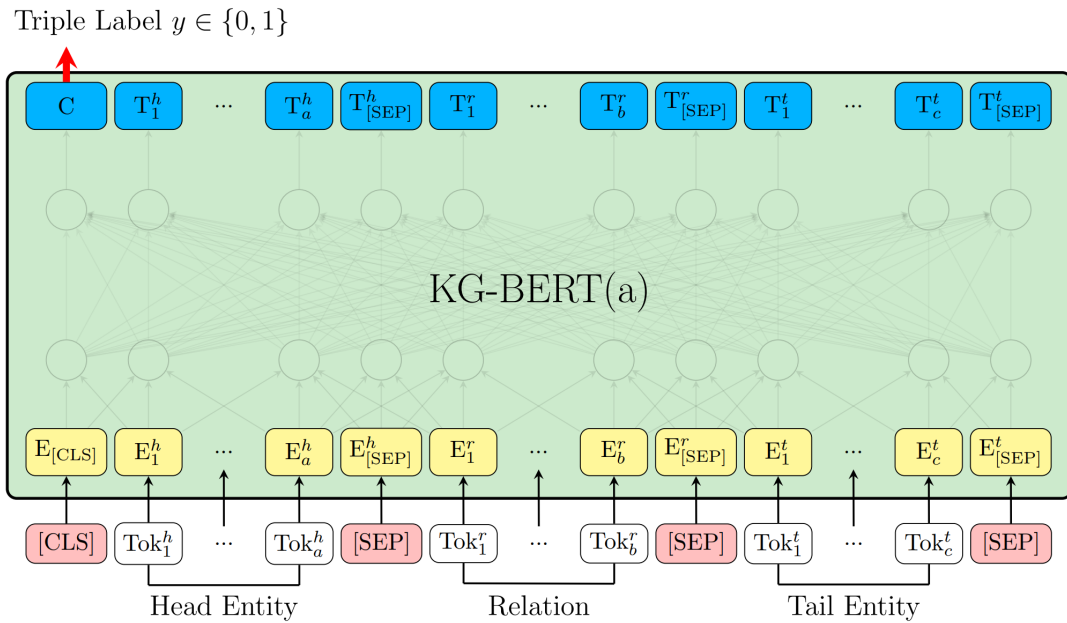


図 1: KG-BERT model [2]

入力にはエンティティの説明文とリレーションの記号「_」を除いたものをいれています。
基本的な BERT のモデルを使用しているようで、隠れ層は 12 層あります。

表 4: 入力例

Entity to text	
14854262	stool, solid excretory product evacuated from the bowels
00590383	chieftainship, the position of chieftain
01990168	founder, sink below the surface
Relation to text	
_member_of_domain_usage	member of domain usage
_has_part	has part

4 to do

4.1 KG-BERT に対してホームズデータセットの適用を検討

ホームズデータセットにはエンティティとリレーションに説明文がないです. WN18RR のエンティティの説明文とホームズデータセットのエンティティを組み合わせるか (エンティティ”Holmes”が WN18RR ないと説明文なしになります.), Wikipedia などからホームズデータセットのエンティティの説明文をもってくるか.

4.2 Relphormer の理解

コード見てがんばります.

参考文献

- [1] Zhen Bi, Siyuan Cheng, Jing Chen, Xiaozhuan Liang, Ningyu Zhang, Qiang Chen, Feiyu Xiong, Wei Guo, and Huajun Chen. Relphormer: Relational graph transformer for knowledge graph representations. 2023.
- [2] Liang Yao, Chengsheng Mao, and Yuan Luo. KG-BERT: BERT for knowledge graph completion. *CoRR*, Vol. abs/1909.03193, , 2019.