

報告書

1 今週の進捗

- テスト数 10, 100, 250, 500 の実験結果の比較
- 訓練データの確認
- エンティティ数 (40,943) 分類の実行

2 KG-BERT [1]

2.1 モデルの説明

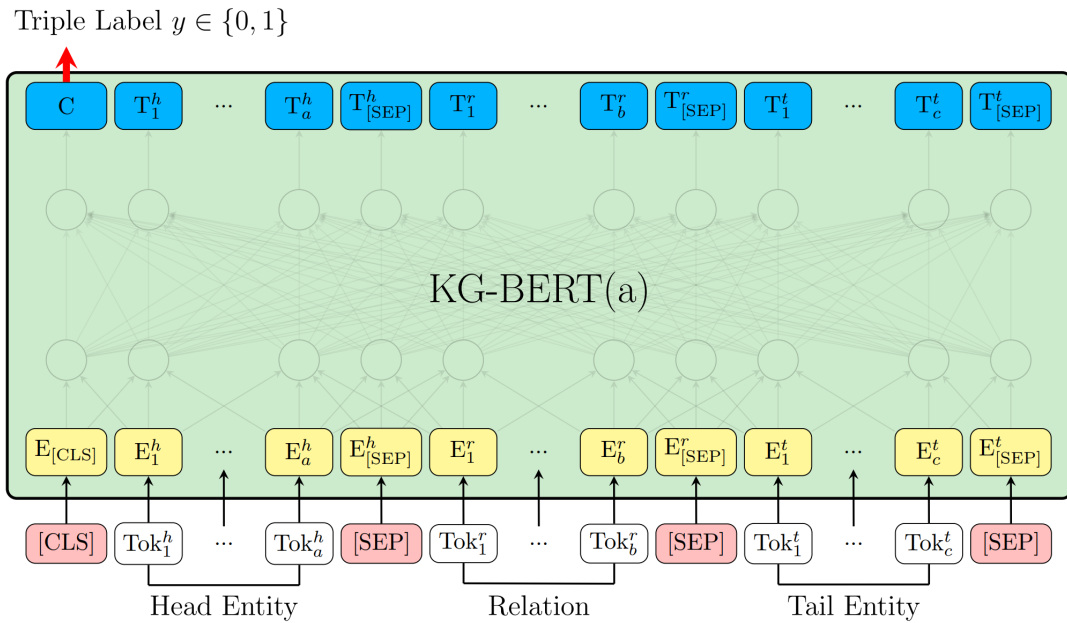


図 1: KG-BERT model [1]

表 1: データセット

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

評価指標として Mean Rank (MR), Mean Reciprocal Rank (MRR), Hits@ k を使用する. MR とは, 予測したエンティティのランクの平均を指す. MRR とは, 予測したエンティティのランクの逆数をスコアとしており, こうして得たスコアの平均をとったものを指す. Hits@ k とは, 予測したエンティティを順位付けしたときに, 上位 k 個以内に正解が含まれている割合のことを指す.

表 2: 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	32
eval batch size	128 (5000)
max seq length	32 (50)

$$\text{MR} = \frac{1}{|E|} \sum_{i=1}^{|E|} \text{rank}_i \quad (1)$$

$$\text{MRR} = \frac{1}{|E|} \sum_{i=1}^{|E|} \frac{1}{\text{rank}_i} \quad (2)$$

$|E|$ はエンティティ数, rank_i は予測したエンティティのランクを表している.

MR は値が小さいとき, MMR, Hits@ k はともに値が大きいくとき推定精度が良いと判断される.

2.2 テスト数 10, 100, 250, 500 の実験結果の比較

テストデータからランダムにテスト数 {10, 100, 250, 500} だけトリプルを取ってきてテストする. 表 3 にそれぞれの結果を示す. テスト数が 10 のときは結果に偏りが生じ, 文献値と大きく異なる結果となった. テスト数が 100 のときまだ途中だが, テスト数が 250, 500 の実験はまだ回せていない.

表 3: テスト数ごとの実験結果

テスト数	WN18RR					実験時間
	MR	MRR	Hits@1	Hits@3	Hits@10	
3,134 (文献値)	97	-	-	-	52.4	
3,134 (再現実験)	117.77	0.25	12.41	29.44	51.85	約 10 日
10	4.5	0.225	0.0	0.0	100.0	約 4 日
100	77.99				53.40	約 5.5 日
250						
500						

2.3 訓練データの確認

1 つの訓練トリプルに対して, ランダムで head か tail かを選択し, 選択した head, tail が不正解となる 5 つのトリプルを生成している. 表 4 にその例を示す.

2.4 エンティティ数 (40,943) 分類

ローカルで実験を回しそうとしています.

表 4: tail が選ばれた場合, 訓練データの不正解トリプルの生成例

	head	relation	tail
正解	land reform, a redistribution of ~	hypernym	reform, a change for ~
不正解	land reform, a redistribution of ~	hypernym	platycerium, often ~
	land reform, a redistribution of ~	hypernym	musculus, one of the ~
	land reform, a redistribution of ~	hypernym	stride, walk with long ~
	land reform, a redistribution of ~	hypernym	upshot, a phenomenon that ~
	land reform, a redistribution of ~	hypernym	attorneyship, the position ~

3 今後したいこと

-

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.