**学习周报**

**2021.8.2 -2021.8.8**

**汇报人：田昊**

1. 本周成果
   * + 1. 知识补全常用数据集、模型、实验任务、评估指标
       2. 传统KG Embedding methods（TransE，TransH，TransR，TransD）
       3. 知识补全和plm的结合论文

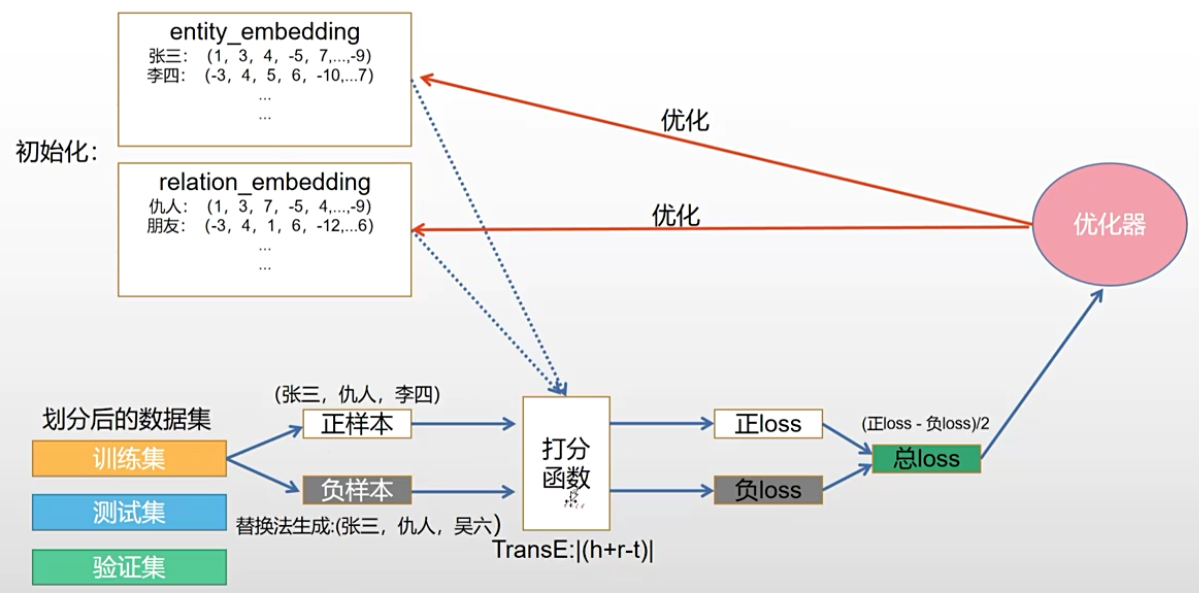
Liang Yao, Chengsheng Mao, Yuan Luo. 2020AAAI.

KG-BERT: BERT for Knowledge Graph Completion.

* + - 1. 运行TransE代码
      2. 运行KG-BERT代码

1. 本周总结
   * + 1. 论文思路：BERT+KGC in Chinese
       2. 知识补全大体思路：

知识表示学习（知识嵌入）🡪链接预测/关系预测/三元组预测🡪得分函数🡪损失函数🡪训练模型



1. 下周计划
   * + 1. 继续学习KG-BERT代码
       2. 学习模型DisMult, NTN, ConvKB, ConvE, DKRL, ComplEx
2. 学习笔记
   * + 1. 知识补全实验所用数据集

* WordNet子集 WN18
* FreeBase子集 FB15k
* WN18RR（去掉数据集 WN18 所有反转关系）
* FB15K-237（去掉数据集 FB15k 所有反转关系）
  + - 1. 知识补全论文实验任务
* 三元组预测Triple Classifification——判断给定的三元组(h,r,t)是否正确
* 链接预测Link Prediction——预测三元组缺失实体(?,r,t)(h,r,?)
* 关系预测Relation Prediction——预测三元组缺失关系(h,?,t)
  + - 1. 知识补全评估指标
* 三元组预测Triple Classifification

准确率、查全率、F1值

* 链接预测Link Prediction/关系预测Relation Prediction

对于每个三元组triple，采用所有的实体替换 triple 中的头尾实体产生corrupted triplets，通过模型计算corrupted triplets的得分，按照升序排名，最终正确triple的排名也确定了。值越小，三元组越逼近正确标准。

Raw Mean Rank（正确的triple在得到的得分序列中的平均排名）

Raw Hits@ k（正确的triple在得分为前 k 的元素中出现的概率）

有其他的正确实体排在待测的正确实体之前的情况（（奥巴马，是总统，美国）替换头结点变成（特朗普，是总统，美国），而（特朗普，是总统，美国）也是已有的正确实体）。

Filtered Mean Rank（删去了其他的正确triple后目标实体的平均排名）

Filtered Hits@ k （删去了其他的正确triple后目标实体在前 k 个元素中出现的概率）

* + - 1. 模型：

**KG embedding methods:**

1. TransE (Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; andYakhnenko, O. 2013. Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems 26. Curran Associates, Inc. 2787–2795.)
2. TransH (Wang, Z.; Zhang, J.; Feng, J.; and Chen, Z. 2014b. Knowledge graph embedding by translating on hyperplanes. In *AAAI*.)
3. TransD (Ji, G.; He, S.; Xu, L.; Liu, K.; and Zhao, J. 2015. Knowledge graph embedding via dynamic mapping matrix. In *ACL*, 687–696.)
4. TransR (Lin, Y.; Liu, Z.; Sun, M.; Liu, Y.; and Zhu, X. 2015b. Learning entity and relation embeddings for knowledge graph completion. In *AAAI*.)
5. TransG (Xiao, H.; Huang, M.; and Zhu, X. 2016. TransG: A generative model for knowledge graph embedding. In *ACL*, volume 1, 2316–2325.)
6. TranSparse (Ji, G.; Liu, K.; He, S.; and Zhao, J. 2016. Knowledge graph completion with adaptive sparse transfer matrix. In *AAAI*.)
7. PTransE (Lin, Y.; Liu, Z.; Luan, H.; Sun, M.; Rao, S.; and Liu, S. 2015a. Modeling relation paths for representation learning of knowledge bases. In *EMNLP*, 705–714.)
8. DistMult and its extension DistMult-HRS which only used structural information in KG. (Zhang, Z.; Zhuang, F.; Qu, M.; Lin, F.; and He, Q. 2018. Knowledge graph embedding with hierarchical relation structure. In *EMNLP*, 3198–3207.)

**The neural tensor network:**

1. NTN (Socher, R.; Chen, D.; Manning, C. D.; and Ng, A. 2013. Reasoning with neural tensor networks for knowledge base completion. In *NIPS*, 926–934.)
2. NTN’s simplifified version ProjE (Shi, B., and Weninger, T. 2017. ProjE: Embedding projection for knowledge graph completion. In *AAAI*.)

**CNN models:**

1. ConvKB (Nguyen, D. Q.; Nguyen, D. Q.; Nguyen, T. D.; and Phung, D. 2018a. A convolutional neural network-based model for knowledge base completion and its application to search personalization. *Semantic Web*.)
2. ConvE (Dettmers, T.; Minervini, P.; Stenetorp, P.; and Riedel, S. 2018. Convolutional 2d knowledge graph embeddings. In *AAAI*, 1811–1818.)
3. R-GCN (Schlichtkrull, M.; Kipf, T. N.; Bloem, P.; Van Den Berg, R.; Titov, I.; and Welling, M. 2018. Modeling relational data with graph convolutional networks. In *ESWC*, 593–607.)

**KG embeddings with textual information:**

1. TEKE (Wang, Z., and Li, J.-Z. 2016. Text-enhanced representation learning for knowledge graph. In *IJCAI*, 1293–1299.)
2. DKRL (Xie, R.; Liu, Z.; Jia, J.; Luan, H.; and Sun, M. 2016. Representation learning of knowledge graphs with entity descriptions. In *AAAI*.)
3. SSP (Xiao, H.; Huang, M.; Meng, L.; and Zhu, X. 2017. SSP: semantic space projection for knowledge graph embedding with text descriptions. In *AAAI*.)
4. AATE (An, B.; Chen, B.; Han, X.; and Sun, L. 2018. Accurate text-enhanced knowledge graph representation learning. In *NAACL*, 745–755.)

**KG embeddings with entity hierarchical types层次类型:**

1. TKRL (Xie, R.; Liu, Z.; and Sun, M. 2016. Representation learning of knowledge graphs with hierarchical types. In *IJCAI*, 2965–2971.)

**Contextualized KG embeddings:**

1. DOLORES (Wang, H.; Kulkarni, V.; and Wang, W. Y. 2018. Dolores: Deep contextualized knowledge graph embeddings. *arXiv preprint arXiv:1811.00147*.)

**Complex-valued KG embeddings:**

1. ComplEx (Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, ´E.; and Bouchard, G. 2016. Complex embeddings for simple link prediction. In *ICML*, 2071–2080.)
2. RotatE (Sun, Z.; Deng, Z.-H.; Nie, J.-Y.; and Tang, J. 2019. Rotate: Knowledge graph embedding by relational rotation in complex space. In *ICLR*.)

**Adversarial learning framework对抗学习框架:**

1. KBGAN (Cai, L., and Wang, W. Y. 2018. KBGAN: Adversarial learning for knowledge graph embeddings. In *NAACL*, 1470–1480.).
   * + 1. 几篇论文实验所用模型

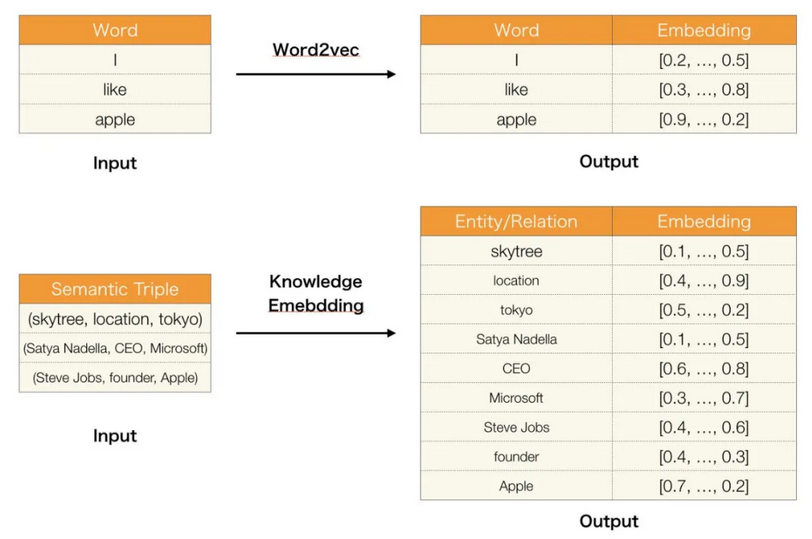
TransE, TransH, TransR, TransD, DisMult, NTN, ConvKB, ConvE, DKRL, ComplEx

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 胶囊网络 | 结构与文本表示 | 实体相似度 | 组合关系路径 | 三元组重要性 | 半边原理 | 卷积网络 |
| 链接预测  （实体预测、关系预测） | DistMult  ComplEx  ConvE  TransE  ConvKB | TransE  TransR  CNN  word2vec( SG) + TransE  CNN + TransE | UM  TransE  TransH  TransR  TranSparse(US)  TranSparse(S)  TransD  TransF  ProjE  LCPE | TransE(unif)  TransR(unif)  TransE(bern)  TransR(bern)  TransH(unif)  TransH(bern)  CRPC(unif)  CRPC(bern) | TransE  TransH  TransR  DisMult  ComplEx  EnTransE  ImpTransE | TransE  DKRL(CBOW)  DKRL(CNN)  DKRL(CNN) + TransE  HMCM | IRN  KBGAN  DisMult  ComplEx  ConvE  TransE  ConvKB |
| 三元组分类 | TransE  TransH  TransR  CTransR | NTN  TransE  TransR  CNN  word2vec( SG) + TransE  CNN + TransE |  |  |  |  | NTN  TransD  TransE DKRL(CBOW)  TransH DKRL(CNN)  TransR  CTransR |

* + - 1. Trans系列模型

知识图谱嵌入/知识表示的转换模型

类似W2V，分布式表示KG里的三元组 (head, relation, tail)



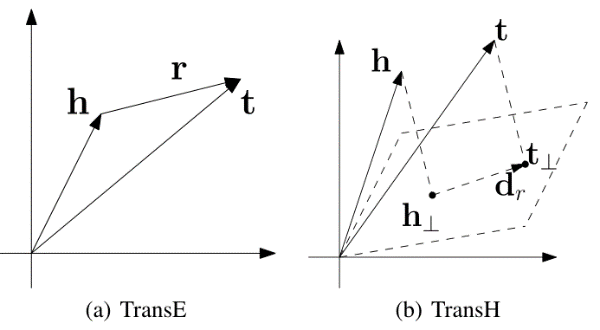
1. TransE模型（E-Embedding）

Title: Translating Embeddings for Modeling Multi-relational Data.

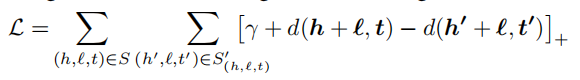
Author：Bordes A, Usunier N, Garciaduran A, et al.

Accepted: Neural Information Processing Systems, NIPS2013.

基本思想就是把 relation 看做是 head 到 tail 的翻译，认为一个正确的知识三元组应该满足 h + r ≈ t，而错误的不满足，即头实体 embedding 加上关系 embedding 近似等于尾实体 embedding



损失函数：



其中距离函数d(h + r, t) = || h + r – t ||，可使用L1或L2范数，S是正确的三元组，S’是通过替换 h 或 t 所得错误的三元组corrupted triplet，γ是间隔距离超参数margin，[x]+表示正值函数，即 x > 0 时，[x]+ = x；当 x ≤ 0 时，[x]+ = 0

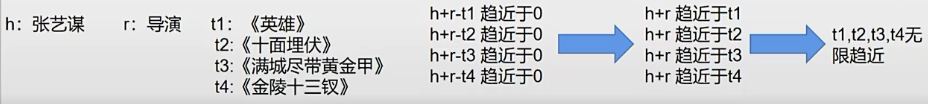
1. TransH模型（H-Hyperlanes）

Title: Knowledge graph embedding by translating on hyperplanes.

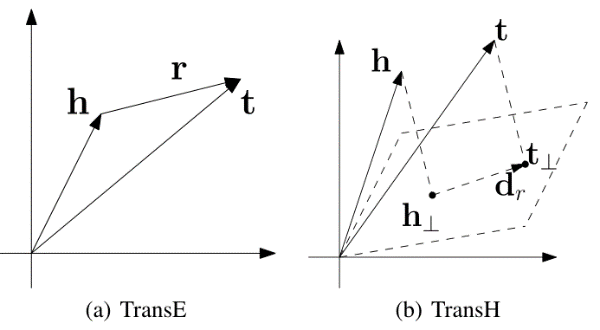
Author：Wang Z, Zhang J, Feng J, et al.

Accepted：National conference on artificial intelligence, AAAI 2014.

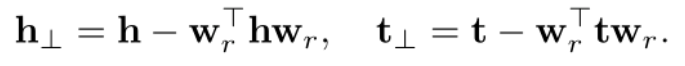
提出原因：TransE 的对简单关系的建模效果显著，但是对复杂关系的建模效果却十分不理想，如1-N，N-1，N-N这样的复杂关系



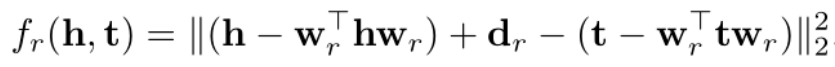
TransH 的核心思想是对于关系每一个 r，有一个超平面 Wr 和一个关系向 dr表示，而不是和实体在同一个嵌入空间。具体来说，在每个三元组中，将头实体 h 和尾实体 t 都映射到这个超平面上得到向量h⊥和t⊥，训练使 h⊥ + dr ≈ t⊥。目的是为了能够使得同一个实体在不同关系中有不同的意义，因为不同的关系有不同的法平面。



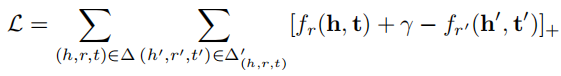
映射关系：



得分函数：



损失函数：



1. TransR模型（R-Relation space）

Title: Learning Entity and Relation Embeddings for Knowledge Graph Completion

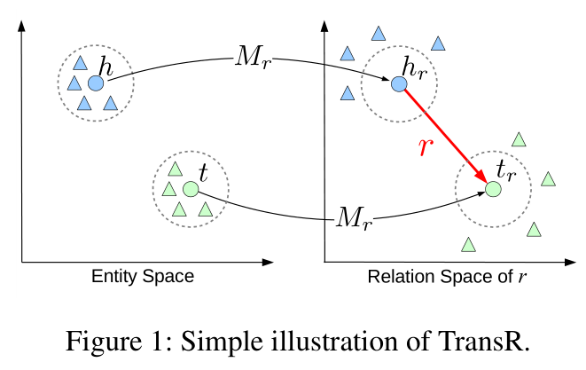
Author：Lin Y, Liu Z, Zhu X, et al.

Accepted：National conference on artificial intelligence, AAAI 2015.

提出原因：TransE，TransH实体和关系是同一语义空间中的向量

TransR的思想是引入两种语义空间，实体空间和关系空间，将头尾实体表示在实体空间，将关系表示在关系空间，这样实体和关系就在不同的语义空间，每种关系r都对应一个关系空间，将头尾实体都映射到关系空间进行训练

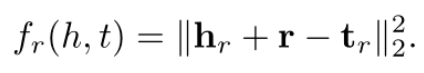
对于每一个三元组 (h, r, t)，对于每一个关系 r，存在一个映射矩阵 Wr，通过这个矩阵将 h, t 映射到关系 r 所在空间，得到 hr 和 tr，使 hr + r = tr。在这种关系的作用下，具有这种关系的头/尾实体（圆圈）彼此接近，不具有此关系的实体（三角形）彼此远离。



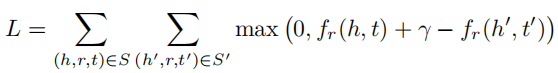
映射关系：



得分函数：



损失函数：



TransR还有一个变体叫做CTransR，其中的C是Cluster聚类。CTransR考虑的问题是一个关系无法体现这一种关系的多义性，比如关系（location，contains，location）其实包含country-city、country-university、continent-country等多种含义。

对于特定的关系r，训练数据中的所有实体对(h，t)聚为多个组，每组中的实体对表现出相似的r关系。所有的实体对(h，t)用它们的向量偏移量(h−t)进行聚类，其中h和t用TransE得到。然后每个聚类的子关系都学习一个变换矩阵Mr和一个表示向量rc。

实体的投影向量为**h**r,c = **hM**r and **t**r,c = **tM**r，得分函数为：



其中旨在确保聚类出的关系向量rc离原始关系向量r不太远，而α控制了该约束的影响。

1. TransD

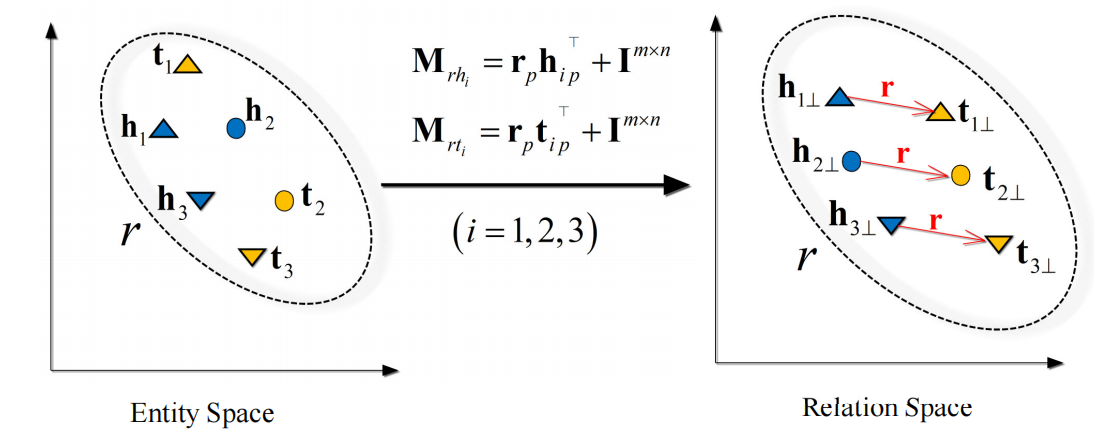
Title: Knowledge graph embedding via dynamic mapping matrix

Author：Ji G, He S, Xu L, et al.

Accepted：ACL 2015

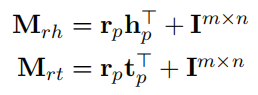
TransR也有其不足之处

* 同一个r，head和tail使用相同的投影矩阵，但是head和tail通常是不同类型的实体，例如，(Bill Gates, founder, Microsoft)。'Bill Gate'是一个人，'Microsoft'是一个公司，这是两个不同的类别。所以他们应该以不同的方式进行转换。
* 投影矩阵仅由关系决定，实际上投影是实体与关系的交互，应与实体和关系都有关。
* TransR的参数数大于TransE和TransH。由于其复杂性，TransR/CTransR难以应用于大规模知识图谱。



TransD方法是将每个实体和关系都由两个向量表示，第一个捕获实体（关系）的含义，另一个投影向量用于构造映射矩阵，这样 (h, r, t) 就有了6个向量**h**, **h**p, **r**, **r**p, **t**, **t**p。

投影矩阵：



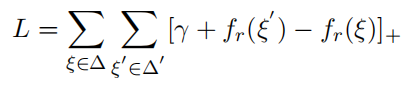
映射关系：



得分函数：



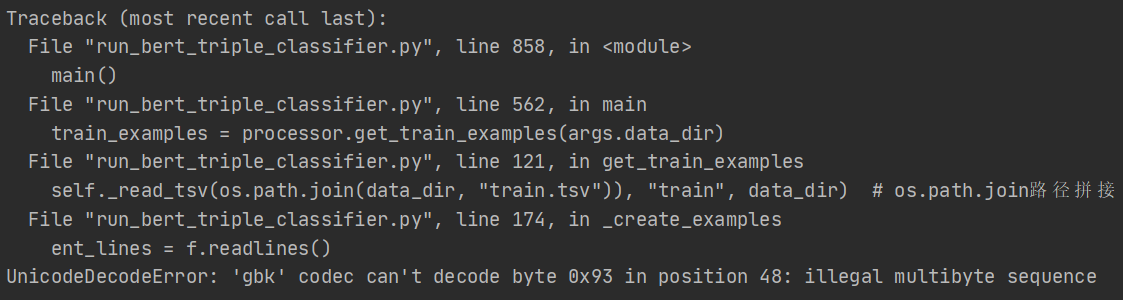
损失函数：



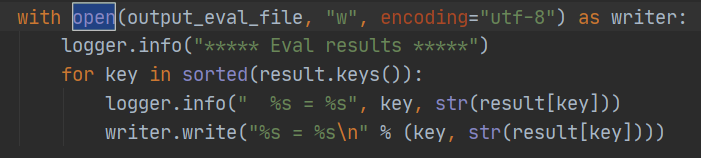
* + - 1. 运行KG-BERT的run\_bert\_triple\_classifier

python run\_bert\_triple\_classifier.py --task\_name kg --do\_train --do\_eval --do\_predict --data\_dir ./data/FB13 --bert\_model bert-base-cased --max\_seq\_length 200 --train\_batch\_size 32 --learning\_rate 5e-5 --num\_train\_epochs 3.0 --output\_dir ./output\_FB13/ --gradient\_accumulation\_steps 1 --eval\_batch\_size 512

出现报错，编码错误



Open函数加上encoding=”utf-8”



导入BERT-BASE-CASED（区分大小写英文）

bert-base-uncased

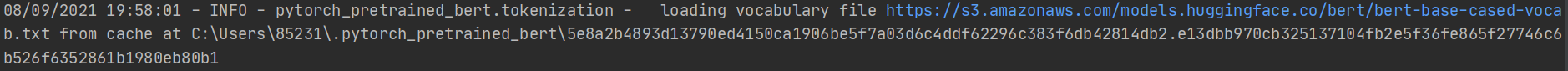
bert-large-uncased

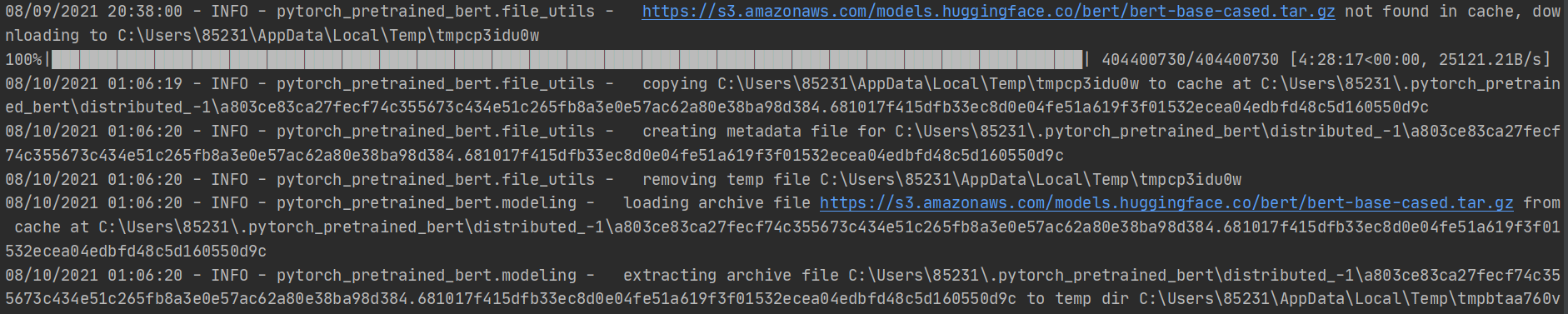
bert-base-cased

bert-large-cased

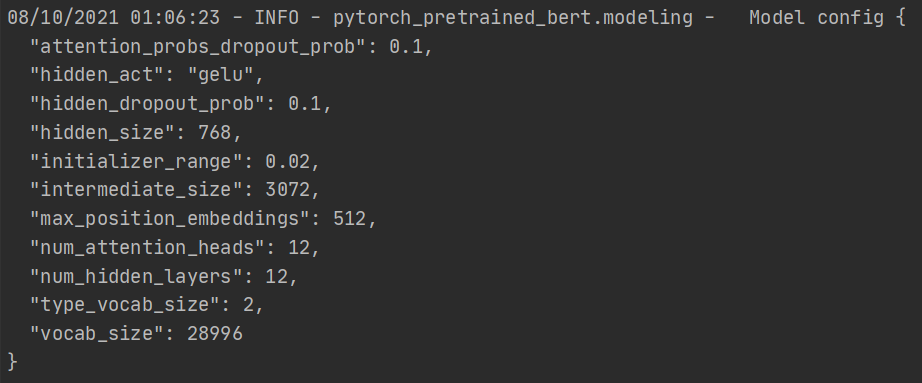
bert-base-multilingual-uncased

bert-base-multilingual-cased, bert-base-chinese

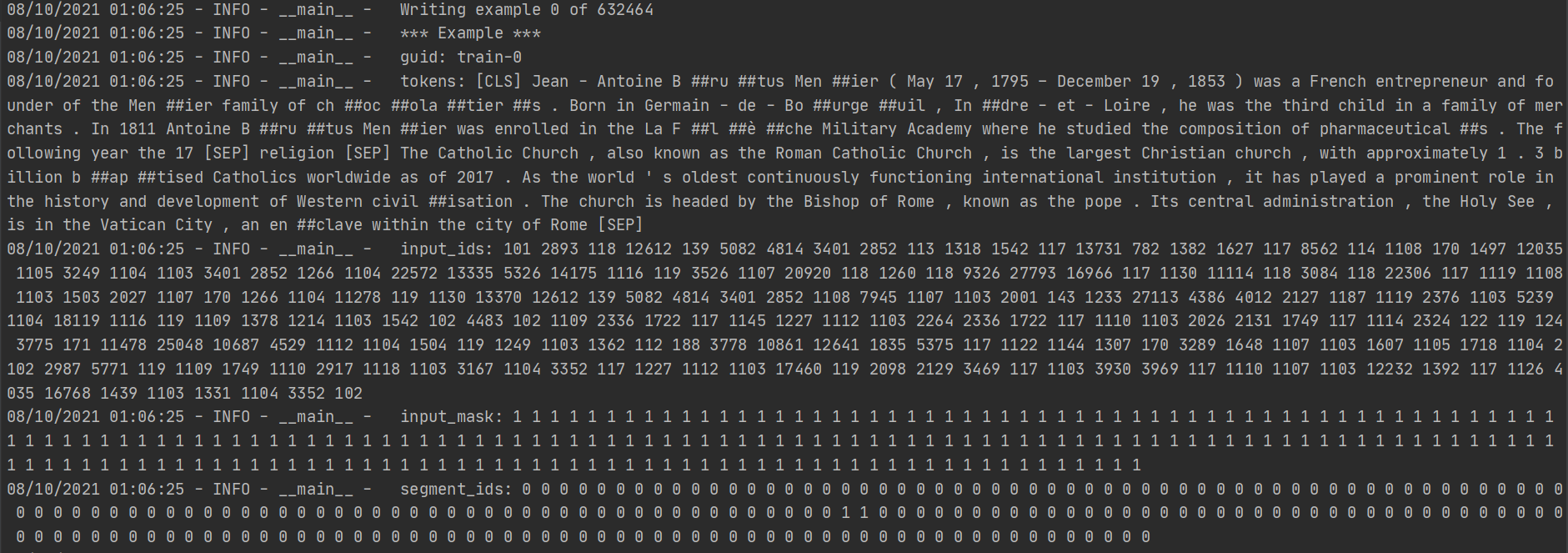


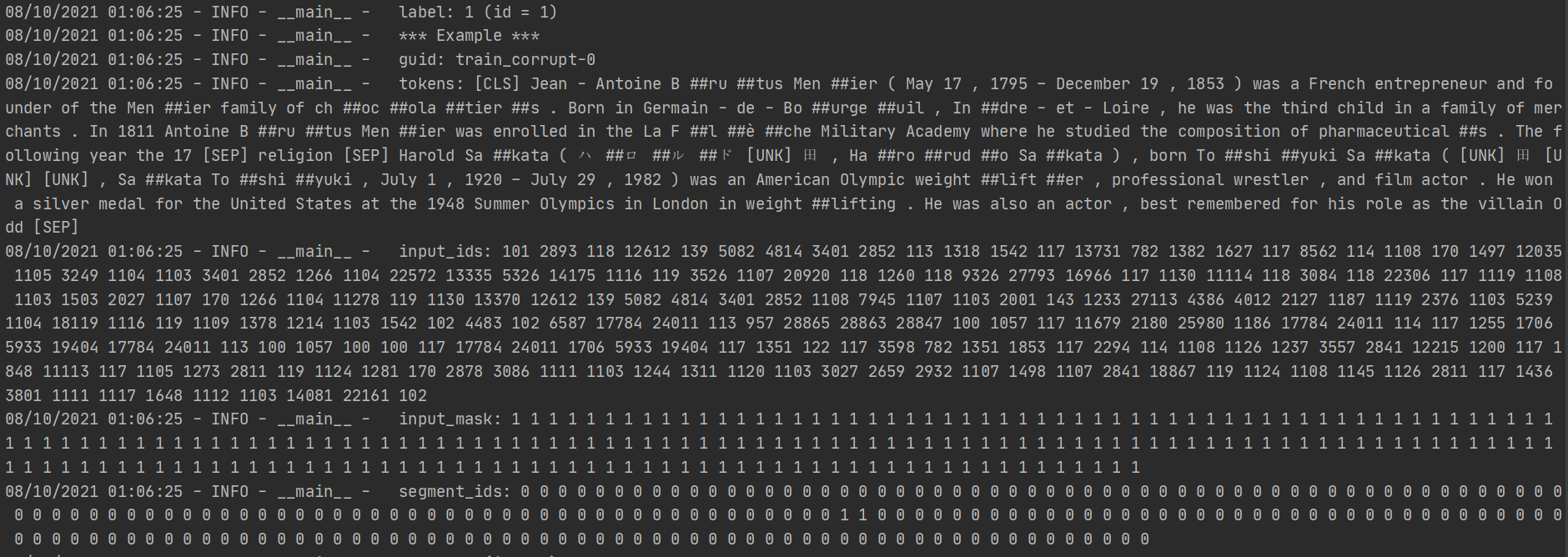


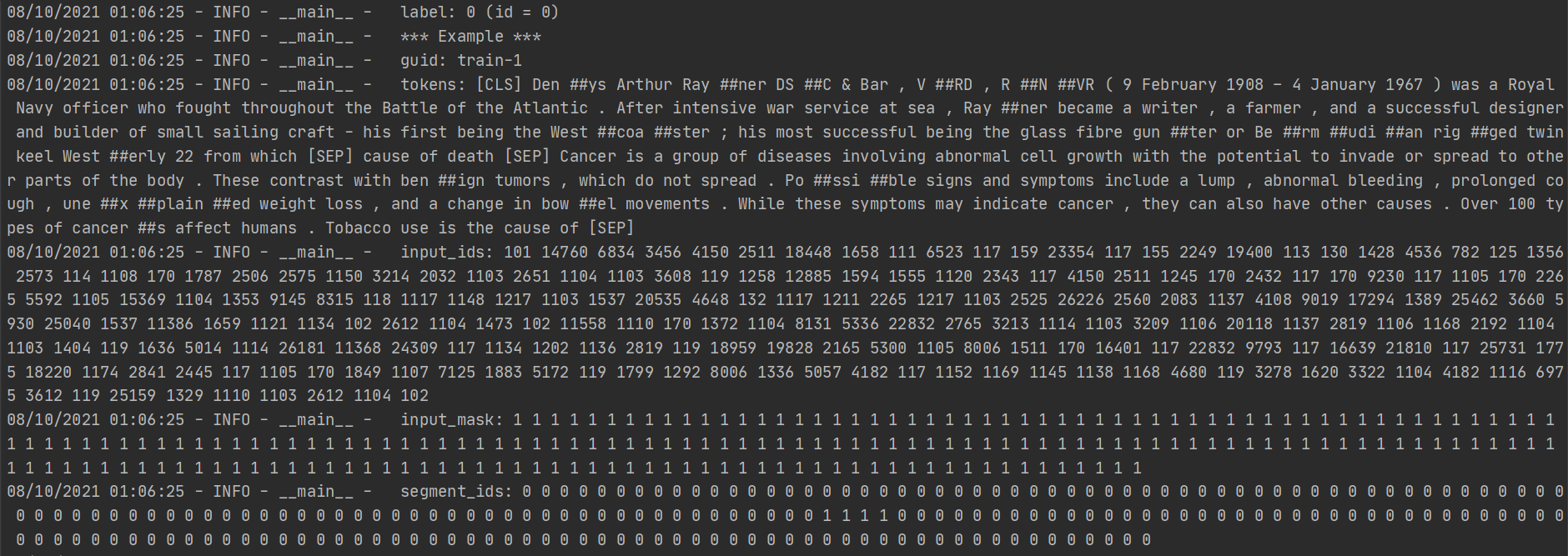
BERT-BASE-CASED参数

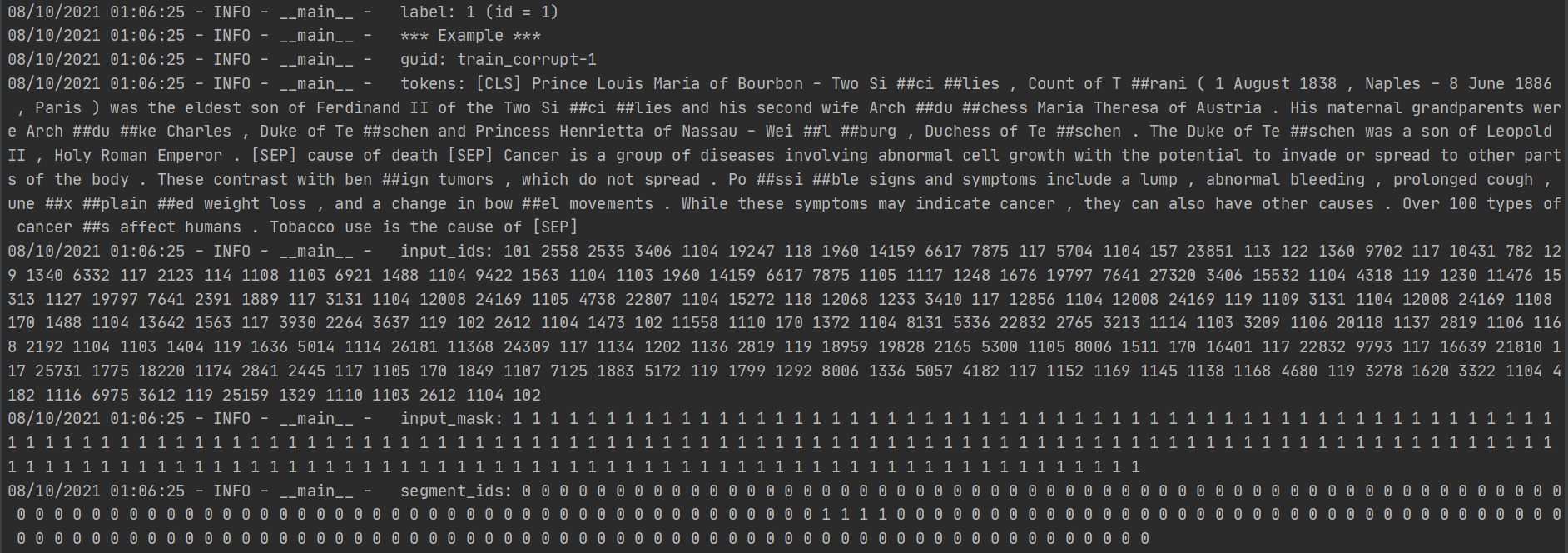


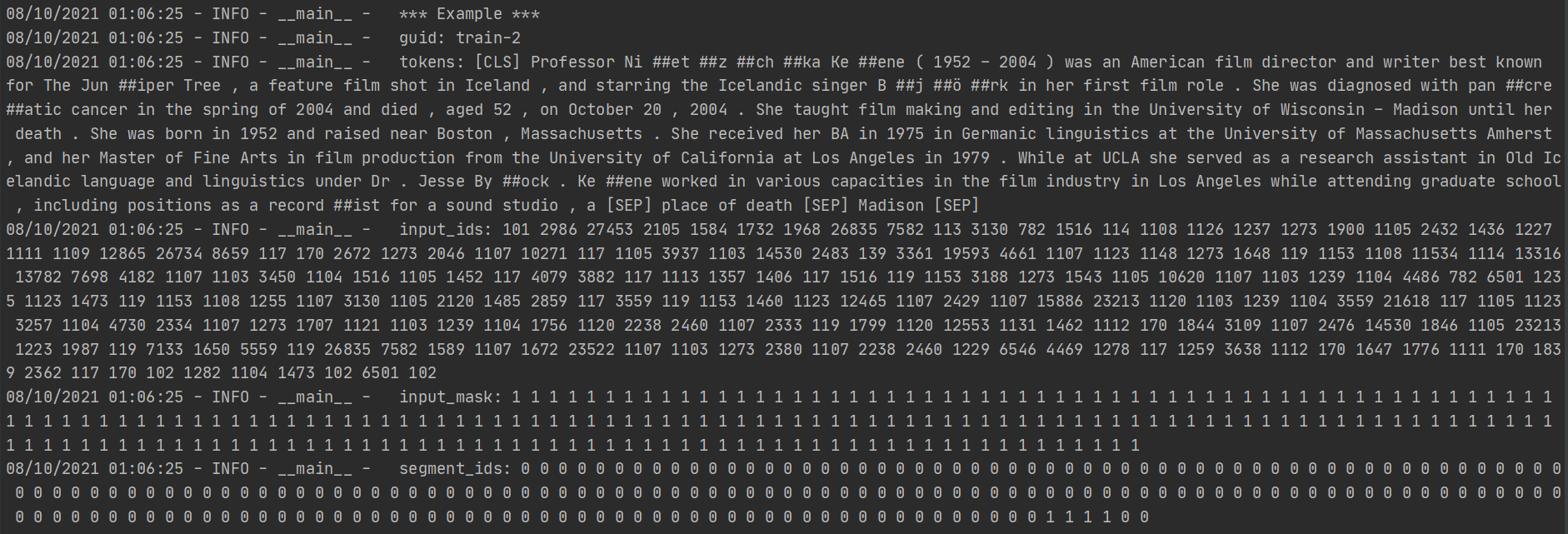
微调？

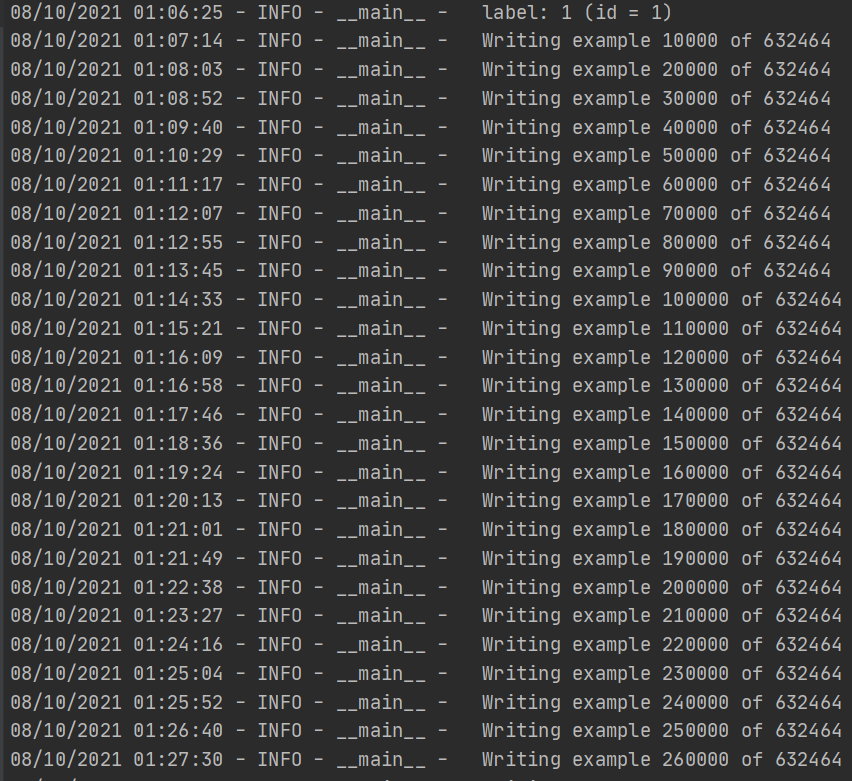




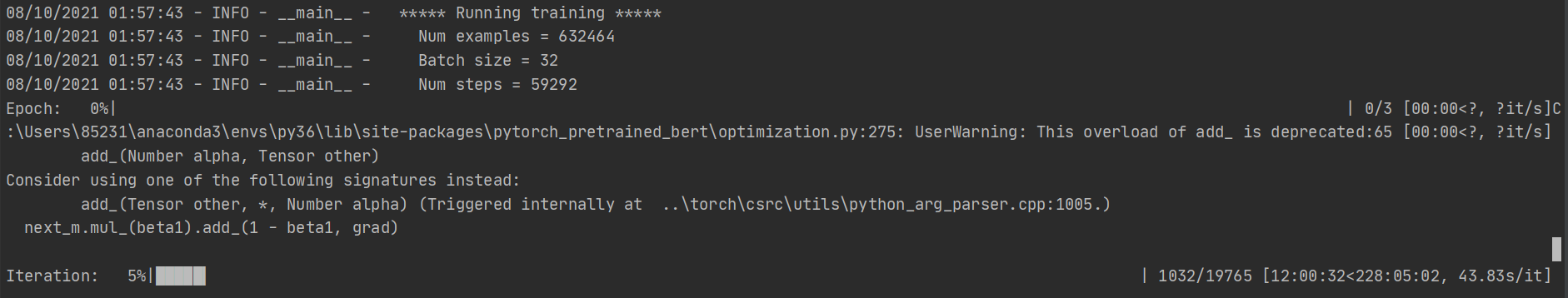








开始训练



组会笔记：

* 综述
* 先中文语料库

再BERT之后加上别的模型再输出CLS

Plm书