

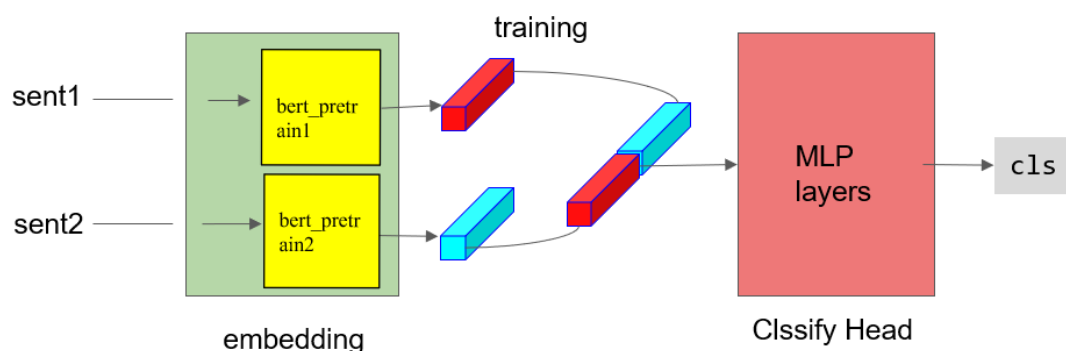
提高OCNLI任务准确率

roberta_large(acc: 74-76?)

尝试架构 Siamese network

使用 Bert_chinese_base:

结构如下



acc 大约50

结果并不理想，有可能是因为学不到推理的语义

论文 StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding

<https://arxiv.org/abs/1908.04577>

ICLR 2020:

结果:

without finetune:

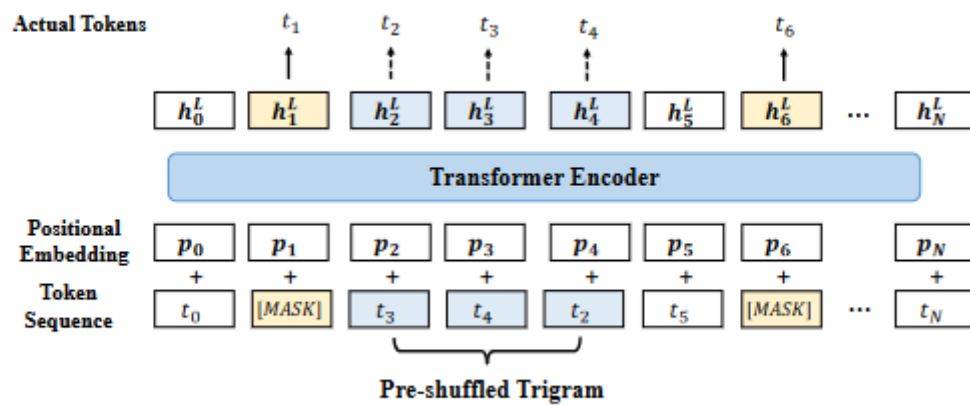
排行	模型	研究机构	测评时间	Score	认证	OCNLI_50K	OC
1	HUMAN-OCNLI	CLUE	20-10-07	90.300	已认证	90.300	{
2	STBert	PCG搜索中台	22-11-25	84.483	待认证	86.933	{
3	ShenZhou	QQ浏览器实验室(QQ Browser Lab)	21-09-19	83.534	待认证	86.367	{
4	Motian	QQ浏览器搜索	21-06-25	82.834	待认证	84.967	{
5	wyd1	wyd1	24-05-10	80.700	待认证	80.700	{

with finetune:

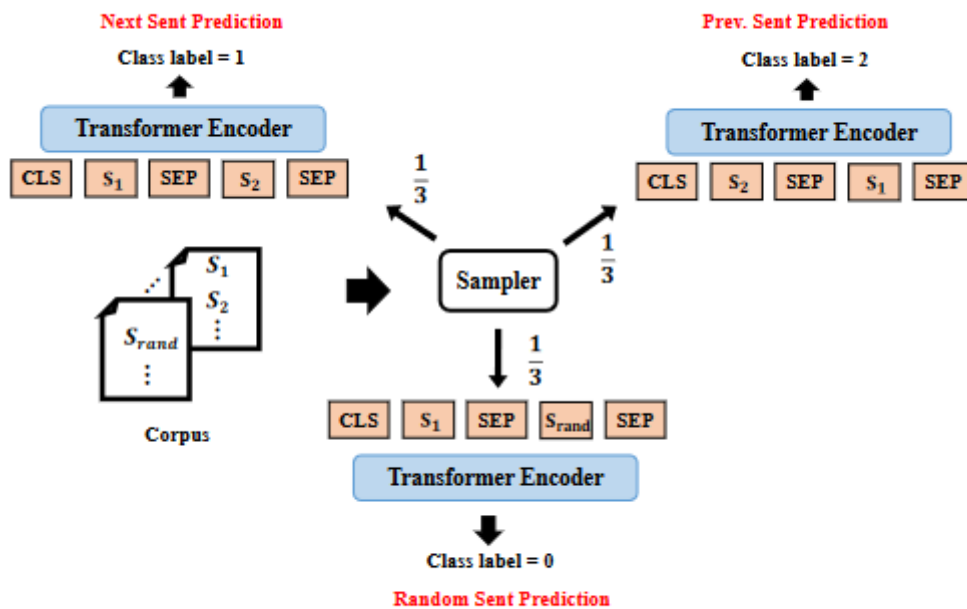
26903	wyd1	24-05-13	11.171	9.775	0	0	0	0	78.2
-------	------	----------	--------	-------	---	---	---	---	------

在损失函数构建时增加任务:

打乱句中单词顺序



(a) Word Structural Objective



(b) Sentence Structural Objective

psp (previous sentence prediction): 预测是否为前句

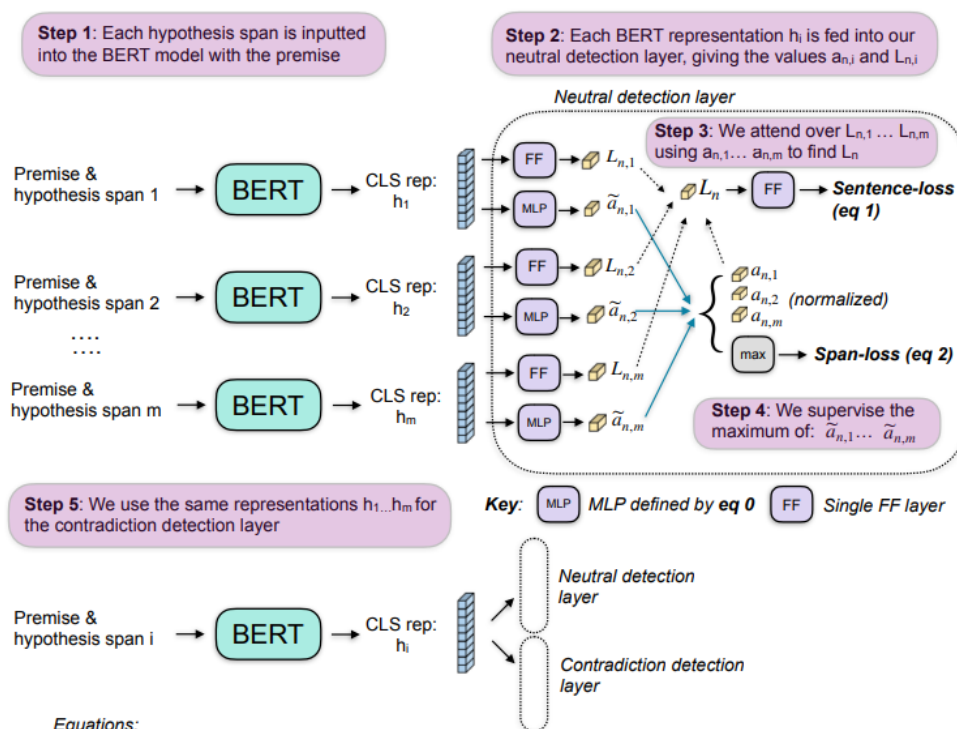
Rsp (random sentence prediction): random \$s_2\$, 查看是否不是连续出现

与nsp区别: 处理较为断裂或无序的文本信息时的鲁棒性, 例如, 在网络抓取的内容或多来源文本聚合时, 内容间可能不具有直接的逻辑关系。

下周任务:

尝试论文

Logical Reasoning with Span-Level Predictions for Interpretable and Robust NLI Models (EMNLP2022)



这些表示随后被送入中立检测层，以计算出与每个跨度相关的值和损失。

输出每个跨度的中立值 ($a_{n,i}$) 和跨度损失 ($L_{n,i}$)

span loss + sequence loss