

论文阅读笔记

Step5

MF1833063, 史鹏, spwannasing@gmail.com

2019 年 6 月 30 日

1 ERNIE: Enhanced Language Representation with Informative Entities

此文章是对bert的一次扩展，提出了知识图谱中的多信息实体（informative entity）可以作为外部知识改善语言表征。

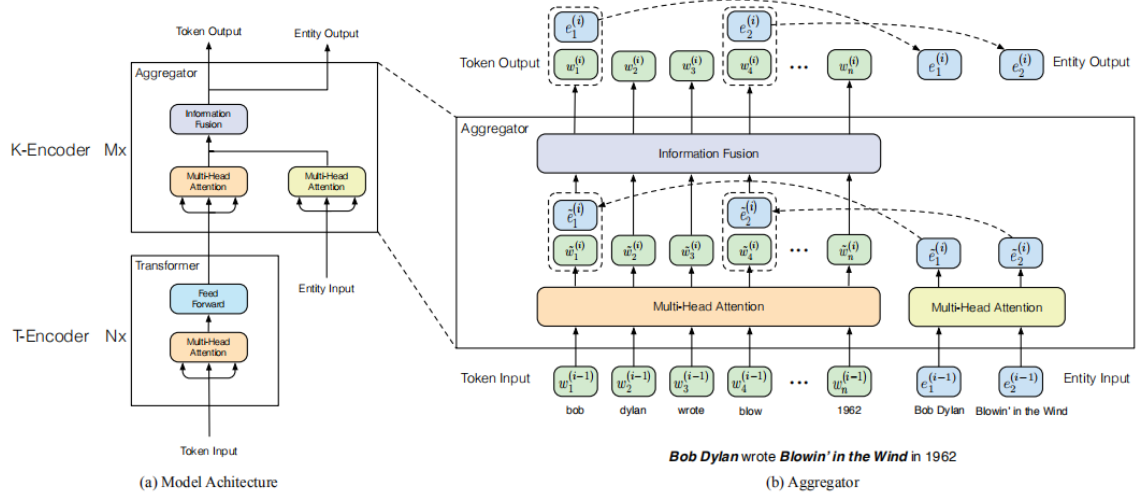


图 1: overview

Mark Twain wrote **The Million Pound Bank Note** in 1893.

Input for Common NLP tasks:

[CLS] [] mark twain [] wrote [] the million pound bank note [] in 1893 . [SEP]

Input for Entity Typing:

[CLS] [ENT] mark twain [ENT] wrote [] the million pound bank note [] in 1893 . [SEP]

Input for Relation Classification:

[CLS] [HD] mark twain [HD] wrote [TL] the million pound bank note [TL] in 1893 . [SEP]

图 2: finetune

Knowledgeable Encoder:

$$\{\tilde{w}_1^{(i)}, \dots, \tilde{w}_n^{(i)}\} = \text{MH-ATT}(\{w_1^{(i-1)}, \dots, w_n^{(i-1)}\}) \quad (1.1)$$

$$\{\tilde{e}_1^{(i)}, \dots, \tilde{e}_m^{(i)}\} = \text{MH-ATT}(\{e_1^{(i-1)}, \dots, e_m^{(i-1)}\}) \quad (1.2)$$

对于和entity对齐的token:

$$\begin{aligned} h_j &= \sigma(\tilde{W}_t^{(i)} \tilde{w}_j^{(i)} + \tilde{W}_e^{(i)} \tilde{e}_k^{(i)} + \tilde{b}^{(i)}) \\ w_j^{(i)} &= \sigma(W_t^{(i)} h_j + b_t^{(i)}) \\ e_k^{(i)} &= \sigma(W_e^{(i)} h_j + b_e^{(i)}) \end{aligned} \quad (1.3)$$

else:

$$\begin{aligned} \mathbf{h}_j &= \sigma \left(\tilde{\mathbf{W}}_t^{(i)} \tilde{\mathbf{w}}_j^{(i)} + \tilde{\mathbf{b}}^{(i)} \right) \\ \mathbf{w}_j^{(i)} &= \sigma \left(\mathbf{W}_t^{(i)} \mathbf{h}_j + \mathbf{b}_t^{(i)} \right) \end{aligned} \quad (1.4)$$

对于引入的信息的pre-training目标:

$$p(e_j | w_i) = \frac{\exp(\text{linear}(\mathbf{w}_i^o) \cdot \mathbf{e}_j)}{\sum_{k=1}^m \exp(\text{linear}(\mathbf{w}_i^o) \cdot \mathbf{e}_k)} \quad (1.5)$$