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。

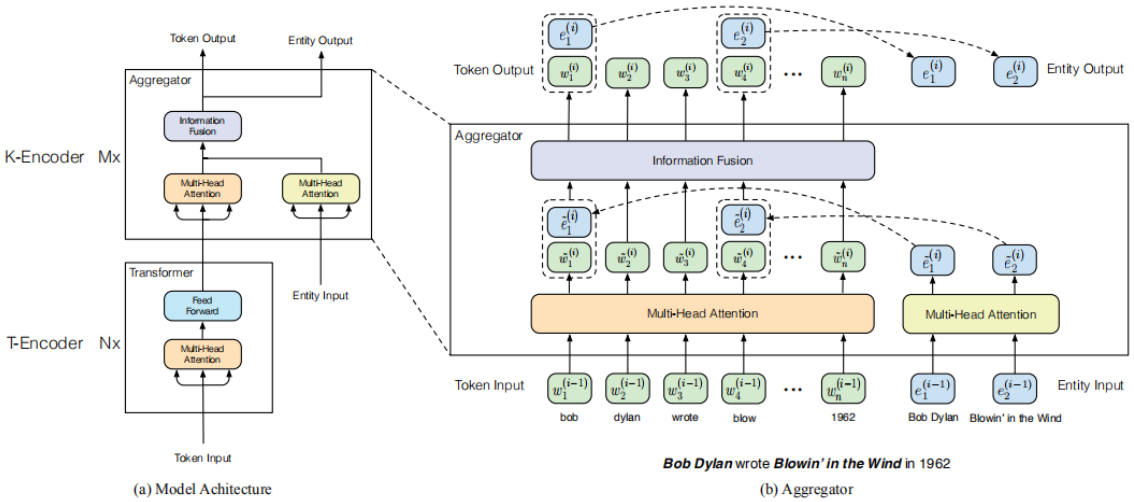

图 1: overview

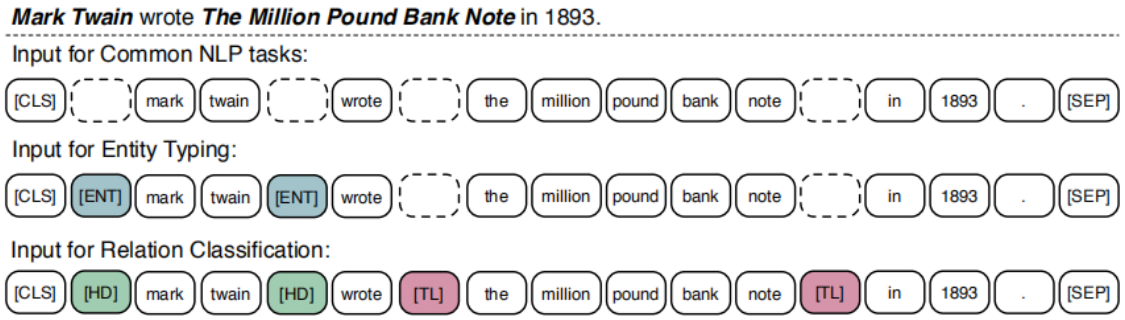

图 2: finetune

Knowledgeable Encoder：

$$\left\{\tilde{\boldsymbol{w}}_1^{(i)}, \ldots, \tilde{\boldsymbol{w}}_n^{(i)}\right\} = \mathrm{MH-ATT}\left(\left\{\boldsymbol{w}_1^{(i-1)}, \ldots, \boldsymbol{w}_n^{(i-1)}\right\}\right) \tag{1.1}$$

$$\left\{\tilde{e}_1^{(i)}, \ldots, \tilde{e}_m^{(i)}\right\} = \mathrm{MH-ATT}\left(\left\{e_1^{(i-1)}, \ldots, e_m^{(i-1)}\right\}\right) \tag{1.2}$$

对于和entity对齐的token：

$$\begin{aligned} \boldsymbol{h}_j &= \sigma\left(\tilde{\boldsymbol{W}}_t^{(i)} \tilde{\boldsymbol{w}}_j^{(i)} + \tilde{\boldsymbol{W}}_e^{(i)} \tilde{\boldsymbol{e}}_k^{(i)} + \tilde{\boldsymbol{b}}^{(i)}\right) \\ \boldsymbol{w}_j^{(i)} &= \sigma\left(\boldsymbol{W}_t^{(i)} \boldsymbol{h}_j + \boldsymbol{b}_t^{(i)}\right) \\ \boldsymbol{e}_k^{(i)} &= \sigma\left(\boldsymbol{W}_e^{(i)} \boldsymbol{h}_j + \boldsymbol{b}_e^{(i)}\right) \end{aligned} \tag{1.3}$$

else：

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

对于引入的信息的pre-training目标：

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