

利用两个自适应特征改进实体链接

张鸿彬, 陈权, 张伟文

广东工业大学计算机学院, 中国广州市, 510006

摘要: 实体链接是自然语言处理中的一项基本任务。现有的基于神经网络的系统更多地关注全局模型的构建, 而忽略了局部模型中潜在的语义信息和有效实体类型信息的获取。本文提出两个自适应特征, 其中第一个自适应特征使得局部和全局模型能够捕获潜在信息, 第二个自适应特征能够描述实体类型嵌入的有效信息。这些自适应特征可以很自然地协同工作来处理一些不确定的实体类型信息。实验结果表明, 我们的实体链接系统在AIDA-B和MSNBC数据集上取得了最佳的性能, 并在域外数据集上达到了最佳的平均性能。这些结果表明, 所提出的自适应特征能够基于其自身不同的上下文来捕获有利于实体链接的信息。

关键词: 实体链接; 局部模型; 全局模型; 自适应特征; 实体类型

<https://doi.org/10.1631/FITEE.2100495>