

观点型问题机器阅读理解中混合词向量和层叠循环神经网络联合训练方法的应用

黄祥洲, 汤斯亮, 张引, 魏宝刚

浙江大学计算机科学与技术学院, 中国杭州市, 310027

摘要: 观点型问题机器阅读理解指计算机通过分析相应段落回答问题。相比于传统机器阅读理解任务的答案是在相关段落中的某一段文本, 观点型问题因其答案可能不出现在相关段落中而需由多个句子推理得出, 其对应的机器阅读理解任务更具挑战性。针对这个任务, 提出一种新颖的基于神经网络的解决方案, 其中使用了一种结合文本特征的混合词向量训练方法。此外, 引入额外的注意力网络和输出层, 产生多个辅助损失函数用于联合训练层叠循环神经网络。针对数据集样本分布不平衡的问题, 引入问题和段落的不相关性实现数据增强。实验结果验证了所提方法的有效性。该方案获得了AIC2018观点型问题机器阅读理解赛道的双周赛冠军。

关键词: 机器阅读理解; 神经网络; 联合训练; 数据增强

<https://doi.org/10.1631/FITEE.1900571>