

# 基于双向注意力语境关联建模的论辩关系预测\*

单华玮, 路冬媛

(对外经济贸易大学 信息学院, 北京 100029)

通信作者: 路冬媛, E-mail: [ludy@uibe.edu.cn](mailto:ludy@uibe.edu.cn)



**摘要:** 在线讨论是当下公众表达意见和交流互动的主要方式之一. 参与者不仅发布评论来表达自己的观点, 还会回复已有的表述进行应答, 支持或反驳他人的观点. 识别表述-应答交互文本的论辩关系可以建模公众对话结构, 挖掘群体意见, 进而为企业产品营销、政府舆情监测等提供辅助. 现有的工作大多通过神经网络对交互文本的条件语义信息或者整体语义信息进行建模, 而忽略了交互文本的语境关联信息. 为此, 提出了一种挖掘语境关联的双向注意力网络模型 (CCRnet). 该模型使用 BERT 分别对表述和应答进行文本语义表示, 并通过双向注意力机制建模交互文本的语境关联表示. 在此基础上, 模型将语境关联表示和交互文本的语义表示进行融合, 建模全局关系特征, 继而输出预测标签. 在 CreateDebate 数据集上的实验结果表明, 与目前主流的方法相比, CCRnet 模型的整体性能表现更优. 此外, 可视化相似度矩阵证实, 双向注意力机制能够有效捕捉交互文本之间的语境关联信息并进一步服务于论辩关系预测.

**关键词:** 机器学习; 深度学习; 论辩关系预测; 注意力机制

**中图法分类号:**

中文引用格式: 单华玮, 路冬媛. 基于双向注意力语境关联建模的论辩关系预测. 软件学报, 2022, 33(5): 1880–1892. <http://www.jos.org.cn/1000-9825/6320.htm>

英文引用格式: Shan HW, Lu DY. Predicting Argumentative Relation with Co-attention Contextual Relevance Network. Ruan Jian Xue Bao/Journal of Software, 2022, 33(5): 1880–1892 (in Chinese). <http://www.jos.org.cn/1000-9825/6320.htm>

## Predicting Argumentative Relation with Co-attention Contextual Relevance Network

SHAN Hua-Wei, LU Dong-Yuan

(School of Information Technology & Management, University of International Business and Economics, Beijing 100029, China)

**Abstract:** Online discussion has become a main way for people to communicate opinions. Besides posting statements, users are also encouraged to reply to existing posts, revealing support or disapproval of others' viewpoints. Identifying argumentative relations between these interactive texts can benefit modeling the dialogue structure, detecting public opinions, and supporting business, marketing, and government to make decisions. Existing studies detected argumentative relations by constructing overall semantic information or conditional semantic information, but the contextual relevance information between interactive texts was ignored. This work proposed a co-attention contextual relevance network (CCRnet). With the co-attention mechanism, the model captured bi-directional attention between the post and reply. Experimental results on the CreateDebate dataset show that the proposed model outperforms the state-of-the-art models. Furthermore, the visualization of the similarity matrix illustrates the effectiveness of the co-attention mechanism.

**Key words:** machine learning; deep learning; argumentative relation prediction; attention mechanism

社交媒体凭借着在内容生成方式、信息传播模式和速度等方面的优势, 快速聚集了庞大的用户群体, 逐渐成为人们日常生活中最主要的信息来源和互动平台. 人们可以随时随地就感兴趣的话题进行在线讨论, 与他人分享

\* 基金项目: 国家自然科学基金面上项目 (62172094); 对外经济贸易大学中央高校基本科研业务费专项资金 (CXTD13-01); 对外经济贸易大学研究生科研创新基金 (202150)

收稿时间: 2020-10-14; 修改时间: 2020-12-01; 采用时间: 2021-02-08; jos 在线出版时间: 2021-08-03

和交流自己的观点. 在讨论过程中, 参与者不仅可以发布评论来表述自己的观点, 还能通过回复对他人的评论进行应答. 人们的应答评论在表达自己观点的同时, 往往还暗含着对他人表述评论的支持或反驳. 如图 1 所示, 在 createdebate 论坛 (<https://www.createdebate.com>) 用户围绕叙利亚难民进行讨论的过程中, 用户 *Cfhsaphg* 发布评论表述了自己的观点, 用户 *Alex\_Marrs* 的应答评论对 *Cfhsaphg* 的观点表达了支持, 而用户 *OptimusLime* 的应答评论则反驳了 *Cfhsaphg* 的观点. 识别表述-应答交互文本的论辩关系可以建模公众讨论的对话结构, 挖掘和分析不同群体的意见, 进而为政府舆情分析和企业品牌营销提供辅助. 因此, 本文面向交互文本进行论辩关系预测研究.

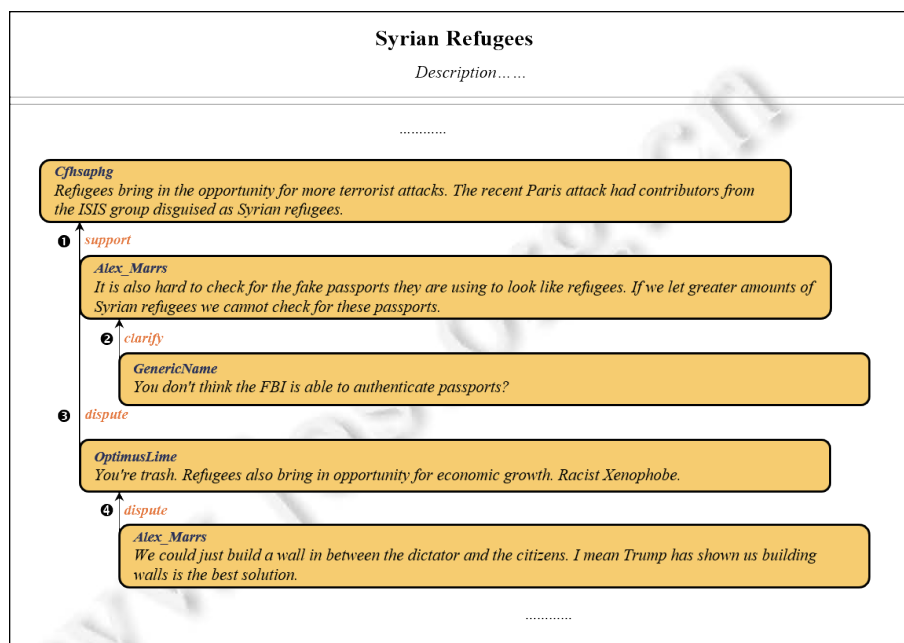


图 1 关于叙利亚难民的在线讨论

现有的论辩关系预测工作面向单文本或者交互文本进行建模. 面向单文本的建模主要关注议论文、法律文书等文本内论点部件 (如: 前提和主张等) 之间的关系. 目前, 基于特征工程的方法和基于深度学习的方法是两种主流的建模方法. 基于特征工程的方法侧重利用词汇、语义等人工构建的特征训练论辩关系分类器<sup>[1-4]</sup>. 基于深度学习的方法则主要通过搭建神经网络进行多任务学习<sup>[5-8]</sup>, 共同解决论点分类和论辩关系预测等任务. 议论文等单文本通常具有完整严谨的论辩结构和直接规范的语言表达, 而具有非结构化和口语化特点的社会媒体文本的质量则参差不齐<sup>[9,10]</sup>. 因而, 面向交互评论文本的论辩关系预测工作更具有挑战性.

传统的面向交互文本的建模方法主要借助主题模型<sup>[11]</sup>和机器学习算法<sup>[12-17]</sup>构建论辩关系预测模型. 这些方法大多使用人工设计的特征建模文本的语义表示, 通常只适用于特定形式的数据集, 普适性较弱. 近年来, 有学者使用基于 LSTM 等深度学习的方法构建论辩关系预测模型, 如: 基于 LSTM 的编码器-解码器模型<sup>[18]</sup>、基于 BERT 的预训练语言模型<sup>[19]</sup>、基于 BiLSTM 的孪生神经网络模型<sup>[20-22]</sup>. 这些方法大多根据交互文本的条件语义信息或者整体语义信息判断两者的论辩关系, 而忽略了交互文本的语境关联信息. 在讨论过程中, 应答文本往往不会直接简明地对表述文本的内容表达支持或反驳. 如图 1 的例句 4 所示, 表述文本和应答文本之间没有任何相同的词汇. 从表面上看, 两者表达的内容似乎是不存在任何关联的. 如果依据表述和应答的整体语义信息或者条件语义信息, 很容易将两者误判为中立关系. 实际上, 应答是在表述内容基础上的回复, 两个文本的词汇在语义上具有一定的关联. 根据表述和应答的交互语境信息, 我们可以直观地判断出应答文本中的“build a wall in between...” (在...之间建墙) 与表述文本中的“Refugees also bring in opportunity for...” (难民也为...带来机会) 是有联系的, 从而帮助模型更准确地识别出两者具有反驳关系.

为此,本文提出了一种挖掘语境关联的双向注意力网络模型 CCRnet (co-attention contextual relevance network),判断表述-应答交互文本之间的支持、反驳或中立关系.具体地,模型使用基于双向 Transformer<sup>[23]</sup>的 BERT<sup>[24]</sup>语言模型对表述和应答分别进行语义表示,并通过双向注意力机制建模表述到应答以及应答到表述两个方向的注意力,得到语境关联表示.在此基础上,模型对表述的文本语义表示和语境关联表示进行融合和池化,得到最受应答关注的表述特征,并将其与应答特征进行拼接,建模全局关系特征,用于输出预测标签.

本文爬取 createdebate.com 辩论论坛的用户评论,构建了 CreateDebate 数据集.为了评估模型的有效性,我们将 CCRnet 模型与目前主流的基线方法和先进方法进行对比.实验结果表明,本文提出的 CCRnet 模型在 CreateDebate 数据集上的整体性能表现优于主流的基线方法和先进方法.

本文的主要贡献:

(1) 提出了一种基于 BERT 和双向注意力机制的孪生神经网络模型,基于孪生神经网络对表述和应答文本同时建模,进行论辩关系预测.

(2) 使用双向注意力机制挖掘表述和应答之间的语境关联表示,通过共享相似度矩阵分别建模表述到应答的注意力和应答到表述的注意力,有效捕捉了两者之间的语境关联表示.

(3) 与多个基线方法和先进方法进行比较,验证了本文提出的 CCRnet 模型以及双向注意力机制的有效性.

本文第 1 节介绍论辩关系预测的相关研究工作.第 2 节详细介绍本文提出的论辩关系预测模型.第 3 节介绍本文的实验设置.第 4 节针对实验结果进行分析.第 5 节总结与展望本文的工作.

## 1 相关工作

论辩关系预测是论辩挖掘 (argument mining) 领域的重要分支之一,旨在识别文本之间的支持、反驳等关系,能够辅助虚假新闻识别<sup>[11]</sup>、群体分类<sup>[15]</sup>等任务.根据建模的角度不同,论辩关系预测可以划分为微观和宏观两种.

微观建模主要面向议论文、法律文本等单文本展开研究,识别单文本内论点部件之间的论辩关系.主流的微观建模方法包括基于特征工程的方法和基于深度学习的方法.早期的基于特征工程的方法通过提取论点部件之间不同方面的特征单独训练论辩关系分类器,如:支持向量机<sup>[1,2]</sup>等.近年来,一些学者提出了对论点分类和论辩关系预测任务进行联合建模的流水线方法<sup>[3,4]</sup>.流水线方法在单独训练每个子任务的基础上,通过整数线性规划 (ILP) 进行全局最优化求解.与流水线方法类似,基于深度学习的方法同样侧重于多任务联合建模,其通过多任务学习的方式<sup>[5-8]</sup>共同解决论点分类和论辩关系预测等任务.多任务学习利用不同子任务的关联信息进行同时训练,避免了流水线方法因单独处理每个子任务产生的错误传播问题.不同于微观建模,本文主要面向社会媒体的在线讨论进行宏观建模,识别表述-应答交互文本之间的论辩关系.

宏观建模主要面向社会媒体的用户生成文本展开研究,识别对话过程中交互文本之间的论辩关系.早期的宏观建模工作应用主题模型<sup>[11]</sup>和机器学习方法<sup>[12-17]</sup>预测论辩关系.其中,有监督的机器学习方法是早期宏观建模的主流方法.例如: Mukherjee 等人<sup>[12]</sup>提出了半监督的 JTE-P 模型挖掘辩论文本的 AD-表达特征,并使用 AD-表达特征训练支持向量机分类器,识别辩论评论之间的赞同或反对关系. Ghosh 等人<sup>[13]</sup>提取文本中的词汇、基于互信息的 unigram 等特征训练支持向量机分类器,识别博客评论之间的支持或反驳关系. Rosenthal 等人<sup>[14]</sup>提取词汇、情感、句子相似度、对话结构等特征训练最大熵分类器.这些方法集中于人工构建和筛选特征,没有充分获取文本的语义信息,停留在浅层语义提取的阶段.

随着深度学习技术的不断发展,其强大的表示学习能力可以自动提取更深层次的语义信息<sup>[25]</sup>.于是,逐渐有学者提出基于 LSTM 等深度学习方法的论辩关系预测模型,解决了对于人工特征的过度依赖问题. Bosc 等人<sup>[18]</sup>提出了基于 LSTM 的编码器-解码器模型,预测推特文本<sup>[26]</sup>之间的支持或反驳关系.该方法通过提取单向的条件语义信息判断论辩关系.然而,模型将编码器末状态的输出作为条件语义向量,无法完整地表示输入文本的语义信息,进而影响解码器的预测效果. Chakrabarty 等人<sup>[19]</sup>受到了 Hewett 等人<sup>[27]</sup>工作的启发,将预训练 BERT 语言模型<sup>[24]</sup>和基于 RST<sup>[28]</sup>的对话关系分类器融合,预测 Reddit/Change My View 板块的交互用户评论之间的论辩关系.虽然模型融合提升了预测效果,但基于 RST 的特征不具有通用性. Bosc 等人<sup>[18]</sup>和 Chakrabarty 等人<sup>[19]</sup>将宏观角度

的论辩关系预测定义为二分类问题. 与之不同, 由于在线讨论存在大量中立关系的交互文本, 本文的论辩关系预测为三分类问题. 除了条件语义模型和预训练语言模型, 有学者从语义匹配的角度构建模型进行论辩关系预测. Cocarascu 等人<sup>[20,21]</sup>提出了基于 BiLSTM 的孪生神经网络模型, 通过拼接或相加的方式匹配交互评论的整体语义表示, 判断交互评论之间的支持、反驳或中立关系. 虽然该方法取得了不错的效果, 但忽视了交互文本的语境关联信息, 无法捕获文本语义表达的焦点, 容易产生语义偏移的问题. 因而, 该方法通常适合处理语义表达直接简明的短文本. Chen 等人<sup>[22]</sup>提出了基于自注意力机制和交叉注意力机制的混合网络模型, 利用两个文本的关键语义信息和语义关联信息判断在线辩论的交互评论之间的支持、反驳或中立关系. 与 Chen 等人<sup>[22]</sup>的工作不同, 本文通过共享权重的相似度矩阵建模表述到应答的注意力表示矩阵和应答到表述的注意力表示矩阵, 让其和表述的文本语义表示进一步进行融合, 减弱了早期加权造成的信息损失.

以往的宏观建模工作大多忽视了表述-应答交互文本的语境关联信息, 挖掘语境关联信息能够帮助模型捕获交互双方论辩的焦点, 避免语义偏移的问题. 为此, 本文提出了一种通过双向注意力机制挖掘语境关联的论辩关系预测模型. 与以往的工作不同, 该模型使用预训练 BERT 语言模型分别编码表述和应答的文本语义表示. 与循环神经网络相比, 预训练 BERT 语言模型具有更强的语义特征表示能力. 此外, 本文通过共享相似度矩阵建模表述和应答之间的语境关联表示, 减弱了早期加权产生的信息损失.

## 2 研究方法

### 2.1 问题定义

面向表述-应答交互文本的论辩关系预测问题的数据集可形式化描述为表述文本、应答文本和论辩关系标签三元组  $\{(P,R),y\}$ . 表述文本  $P$  和应答文本  $R$  分别表示为单词序列, 即  $P = \{p_1, p_2, \dots, p_{L_P}\}$ ,  $R = \{r_1, r_2, \dots, r_{L_R}\}$ . 其中,  $L_P$  和  $L_R$  分别表示表述文本的序列长度和应答文本的序列长度. 表述-应答交互文本的论辩关系标签  $y \in \{1, -1, 0\}$ . 其中,  $y=1$  代表表述和应答之间为支持关系,  $y=-1$  代表反驳关系,  $y=0$  代表中立关系.

### 2.2 模型框架

本文提出了一种挖掘语境关联的双向注意力网络模型 CCRnet, 预测表述-应答交互文本之间的支持、反驳或中立关系. 如图 2 所示, CCRnet 模型包括语义表示层、语义推理层、语义融合层和输出分类层. 语义表示层使用基于双向 Transformer<sup>[23]</sup> 的 BERT 语言模型<sup>[24]</sup> 获取表述和应答的文本语义表示. 语义推理层通过双向注意力机制分别建模表述到应答以及应答到表述两个方向的注意力, 得到语境关联表示. 语义融合层首先通过拼接和池化对表述的文本语义表示和语境关联表示进行编码, 得到最受应答关注的表述特征. 然后, 将其与经过非线性变换和池化后的应答语义表示进行拼接, 建模全局关系特征. 输出分类层通过两层全连接网络输出预测的论辩关系标签.

#### 2.2.1 语义表示层

语义表示层使用共享权重的预训练 BERT 语言模型<sup>[24]</sup> 对表述和应答分别进行语义表示. 近年来, 以 BERT 为代表的预训练语言模型在自然语言处理领域的各类任务中展现了优异的性能. 与传统的 LSTM<sup>[29]</sup>、CNN<sup>[30]</sup> 相比, 基于双向 Transformer<sup>[23]</sup> 的 BERT 模型具有更强的上下文语义特征提取能力. 具体地, 给定表述文本  $P$  和应答文本  $R$ , 语义表示层首先通过词编码和位置编码分别对  $P$  和  $R$  进行嵌入表示. 以表述文本  $P$  为例, 词编码通过词嵌入矩阵将表述中的词  $\{p_i\}_{i=1}^{L_P}$  全部映射为词向量, 得到词嵌入表示  $CE^P \in \mathbb{R}^{d \times L_P}$ . 位置编码通过位置嵌入矩阵将表述中词  $\{p_i\}_{i=1}^{L_P}$  所在的位置全部映射为位置向量, 得到位置嵌入表示  $QE^P \in \mathbb{R}^{d \times L_P}$ . 之后, 将词嵌入表示  $CE^P$  和位置嵌入表示  $QE^P$  相加, 得到最终的表述嵌入表示  $E^P \in \mathbb{R}^{d \times L_P}$ . 同样地, 应答文本  $R$  经过词编码和位置编码, 得到最终的应答嵌入表示  $E^R \in \mathbb{R}^{d \times L_R}$ . 在嵌入表示的基础上, 将  $E^P$  和  $E^R$  分别输入包含 12 层双向 Transformer 的编码层, 得到表述和应答的语义表示矩阵  $M^P$  和  $M^R$ . 其中,  $M^P = [m_1^P, m_2^P, \dots, m_{L_P}^P] \in \mathbb{R}^{d \times L_P}$  和  $M^R = [m_1^R, m_2^R, \dots, m_{L_R}^R] \in \mathbb{R}^{d \times L_R}$ .

#### 2.2.2 语义推理层

在阅读理解<sup>[31]</sup>、立场识别<sup>[32,33]</sup>、文本蕴含<sup>[34]</sup>等研究任务中, 基于注意力机制的模型取得了突破性的进展. 其中, Seo 等人<sup>[31]</sup>使用双向注意力机制对原文和问题之间的关联信息进行建模. 在此基础上, 将两者的关联信息和原文的语义表示进行融合, 得到与问题相关的原文表示, 从而帮助模型明确答案在原文中的位置. 不同于以往的单向

注意力机制, 双向注意力机制通过双向查询的方式捕获了与问题相关的原文信息以及与原文相关的问题信息, 改进了单向注意力机制在提取两个文本的交互信息方面存在的不足. 此外, 双向注意力机制没有直接将原文到问题的注意力和问题到原文的注意力编码为固定长度的特征向量, 而是通过共享相似度矩阵分别计算原文和问题的每个时间步长的注意力权重向量, 使其和原文的文本表征流动到后面的建模层, 减弱了过早加权产生的信息损失.

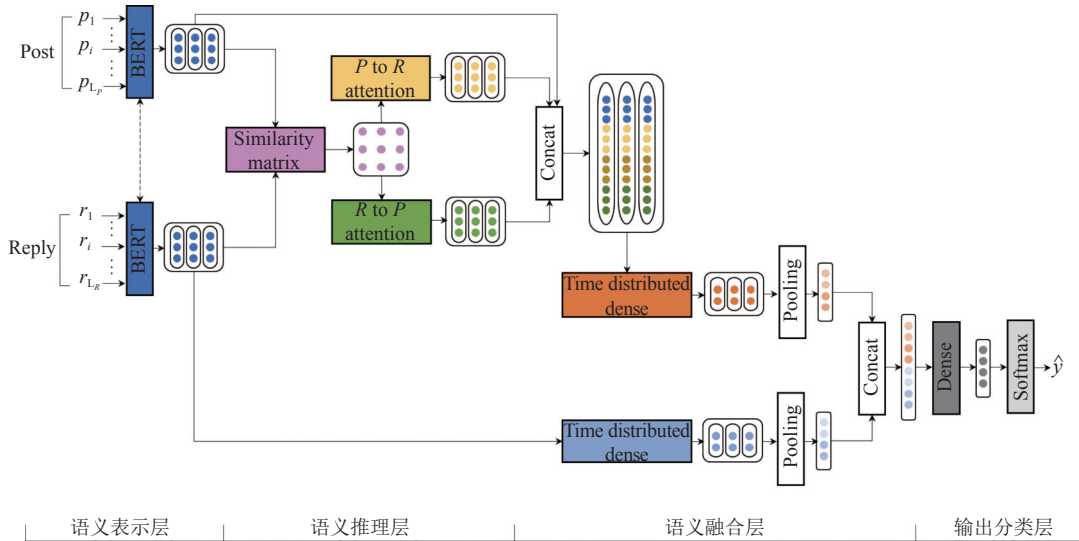


图 2 论辩关系预测模型框架图

本文受到 Seo 等人<sup>[31]</sup>工作的启发, 首次将双向注意力机制引入面向交互文本的论辩关系预测任务, 通过双向注意力机制从表述文本和应答文本的交互语境中捕捉关联信息, 从而帮助模型更好地聚焦交互双方论辩的重点. 双向注意力机制的具体结构如图 3 所示.

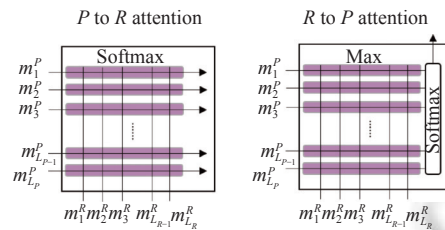


图 3 双向注意力机制的结构图

为了获取表述和应答之间的语境交互信息, 我们通过线性加权计算的方式得到表述与应答的相似度矩阵  $S \in \mathbf{R}^{L_p \times L_r}$ . 它表示表述文本中的每一个词与应答文本中的每一个词之间的相关程度. 具体的计算公式如下所示:

$$S_{ij} = w_s^T [m_i^r; m_j^p; m_i^p \circ m_j^r] \quad (1)$$

其中,  $m_i^p \in \mathbf{R}^d$  表示表述中第  $t$  个词  $p_t$  对应的语义表示向量,  $m_j^r \in \mathbf{R}^d$  表示应答中第  $j$  个词  $r_j$  对应的语义表示向量,  $\circ$  表示两个向量对应位置的元素相乘,  $;$  表示对向量进行拼接,  $w_s \in \mathbf{R}^{3d}$  为可训练的参数,  $S_{ij}$  表示  $m_i^p$  和  $m_j^r$  的相似度值.

根据得到的相似度矩阵  $S$ , 分别建模表述到应答的注意力和应答到表述的注意力, 得到语境关联表示.

(1) 表述到应答的注意力 (P to R attention)

根据相似度矩阵  $S$ , 按行使用 *softmax*, 得到表述中的每一个词和应答中的每一个词之间的注意力权重. 在此基础上, 通过加权求和得到与表述中每一个词相关的应答表示, 从而对应答中与表述相关的信息给予更高的关注度. 最后, 将与表述中每个词相关的应答表示进行拼接, 得到表述到应答的注意力表示矩阵  $\tilde{M}^R \in \mathbf{R}^{d \times L_p}$ . 表述到应

答的注意力捕获了与表述中的每一个词相关联的应答信息. 具体的计算公式如下所示:

$$a_t = \text{softmax}(S_t) \in \mathbb{R}^{L_R} \quad (2)$$

$$\tilde{m}_t^R = \sum_j a_{tj} m_j^R \in \mathbb{R}^d \quad (3)$$

$$\tilde{M}^R = [\tilde{m}_1^R, \dots, \tilde{m}_t^R, \dots, \tilde{m}_{L_P}^R] \in \mathbb{R}^{d \times L_P} \quad (4)$$

其中,  $S_t \in \mathbb{R}^{L_R}$  表示相似度矩阵中第  $t$  行对应的向量,  $a_t$  表示表述中第  $t$  个词对应答中词的注意力权重.  $\tilde{m}_t^R$  表示经过加权求和得到的与表述中第  $t$  个词相关的应答表示.

#### (2) 应答到表述的注意力 (*R to P attention*)

为了获取与应答整体最相关的表述信息, 我们对相似度矩阵  $S$  逐行取最大值, 并通过 *softmax* 得到应答对表述的注意力权重. 在此基础上, 通过加权求和得到与应答相关的表述表示  $\tilde{m}^P$ , 从而对表述中与应答相关的信息给予更高的关注度. 最后, 将  $\tilde{m}^P$  平铺  $L_P$  次得到应答到表述的注意力表示矩阵  $\tilde{M}^P \in \mathbb{R}^{d \times L_P}$ . 应答到表述的注意力捕获了表述中与应答相关联的最重要的词信息. 具体的计算公式如下所示:

$$b = \text{softmax}(\max_{\text{col}}(S)) \in \mathbb{R}^{L_P} \quad (5)$$

$$\tilde{m}^P = \sum_t b_t m_t^P \in \mathbb{R}^d \quad (6)$$

$$\tilde{M}^P = [\tilde{m}^P, \dots, \tilde{m}^P, \dots, \tilde{m}^P] \in \mathbb{R}^{d \times L_P} \quad (7)$$

其中,  $\max_{\text{col}}$  表示逐行取最大值,  $b$  表示应答文本整体对表述中每一个词的注意力权重.

### 2.2.3 语义融合层

在获取表述和应答之间的关联信息后, 我们首先将表述的文本语义表示  $M^P$ 、表述到应答的注意力表示  $\tilde{M}^R$  和应答到表述的注意力表示  $\tilde{M}^P$  进行融合, 得到具有应答感知的表述表示  $C$ . 之后, 通过平均值池化和最大值池化对  $C$  进行特征提取, 并将池化后得到的特征向量进行拼接, 得到最受应答关注的表述特征  $f$ . 具体公式如下所示:

$$C_t = \text{relu}(W_C^T [m_t^P; \tilde{m}_t^R; m_t^P \circ \tilde{m}_t^R; m_t^P \circ \tilde{m}^P] + b_C) \in \mathbb{R}^h \quad (8)$$

$$f = [\text{avgpool}(C); \text{maxpool}(C)] \in \mathbb{R}^{2h} \quad (9)$$

其中,  $W_C \in \mathbb{R}^{4d \times h}$  为可训练的参数,  $b_C$  为偏置项,  $C_t$  表示表述中第  $t$  个词的应答感知表示向量, *avgpool* 表示平均值池化, *maxpool* 表示最大值池化.

为了提取更高层次的应答语义特征, 我们对应答的文本语义表示  $M^R$  进行非线性变换和池化, 并将池化后得到的特征向量进行拼接, 得到应答特征  $z$ . 具体公式如下所示:

$$Q_j = \text{relu}(W_Q^T m_j^R + b_Q) \in \mathbb{R}^e \quad (10)$$

$$z = [\text{avgpool}(Q); \text{maxpool}(Q)] \in \mathbb{R}^{2e} \quad (11)$$

其中,  $W_Q \in \mathbb{R}^{d \times e}$  为可训练的参数,  $b_Q$  为偏置项,  $Q_j$  表示经过非线性变换后的  $m_j^R$  的特征表示.

最后, 将最受应答关注的表述特征  $f$  和应答特征  $z$  进行拼接, 建模全局关系特征  $g$ . 具体公式如下所示:

$$g = [f; z] \in \mathbb{R}^{2h+2e} \quad (12)$$

### 2.2.4 输出分类层

在得到表述-应答交互文本的全局关系特征  $g$  后, 通过输出分类层得到两者的预测类别标签的概率分布  $\hat{y}$ . 该层包含两个前馈神经网络, 第 1 个使用 *tanh* 作为激活函数, 第 2 个使用 *softmax* 作为激活函数. 具体公式如下所示:

$$o = \text{tanh}(W_o^T g + b_o) \in \mathbb{R}^u \quad (13)$$

$$\hat{y} = \text{softmax}(W_y^T o + b_y) \in \mathbb{R}^K \quad (14)$$

其中,  $W_o \in \mathbb{R}^{(2h+2e) \times u}$  和  $W_y \in \mathbb{R}^{u \times K}$  分别为两个前馈神经网络的训练参数,  $b_o$  和  $b_y$  为偏置项,  $K$  表示论辩关系标签的类别总数.

### 2.2.5 目标函数

由于本文的论辩关系预测为三分类问题, 我们采用交叉熵作为损失函数, 用于训练优化模型. 给定训练数据集

$\{(P_i, R_i), y_i\}$ , 交叉熵的计算公式如下所示:

$$loss = - \sum_i \sum_k y_i^k \log \hat{y}_i^k \quad (15)$$

其中, 如果第  $i$  个表述-应答交互文本的真实标签为第  $k$  个类别,  $y_i^k = 1$ , 否则,  $y_i^k = 0$ .  $\hat{y}_i^k$  表示第  $i$  个表述-应答交互文本的预测输出属于第  $k$  个类别标签的概率.

### 3 实验设置

#### 3.1 数据集

本文使用从在线辩论论坛 [createdebate.com](https://www.createdebate.com) 构建的 CreateDebate 数据集评估 CCRnet 模型. 该论坛的用户可以选择任意感兴趣的辩题参与讨论. 在讨论过程中, 参与者不仅可以发布评论来表达观点, 还可以通过回复已有的用户评论与其他用户进行交互, 支持、反驳其他用户的观点. 我们从 17 628 个热门辩题中收集具有支持 (Support)、反驳 (Dispute) 和中立 (Clarify) 关系的表述-应答交互文本, 得到了由 172 259 条带有论辩关系标签的表述-应答交互文本构成的 CreateDebate 数据集. 由于交互文本所属的辩题类型差异性较大且缺乏有效信息, 我们没有额外收集具体的辩题数据. 我们将 CreateDebate 数据集公开发布在 [github](https://github.com/chachashw/createdebate) (<https://github.com/chachashw/createdebate>) 上. 数据集的标签分布如表 1 所示.

表 1 CreateDebate 数据集的标签分布

类别标签	数量
支持	60 384
反驳	67 336
中立	44 539

#### 3.2 模型训练与参数设置

本文按照 8:1:1 的比例将 CreateDebate 数据集划分为训练集、验证集和测试集. 其中, 验证集用于超参数调整, 测试集用于模型性能评估. CCRnet 模型的 BERT 语义表示网络使用 Google 开源的 BERT<sub>Base</sub> 预训练语言模型. 该模型包含 12 层双向 Transformer, 隐层大小为 768 维. 如图 4 的小提琴图所示, 在 CreateDebate 数据集中, 绝大多数的表述和应答均为短文本. 经过统计分析发现, 83.73% 的表述文本的长度低于 100, 86.27% 的应答文本的长度低于 100, 故设定 100 为最大文本长度. 在训练过程中, 我们使用 Adam<sup>[35]</sup>更新模型参数. 模型的超参设置如表 2 所示.

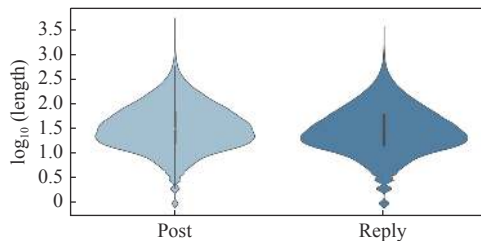


图 4 表述文本和应答文本的长度分布图

表 2 超参数设置表

超参数	数值
学习率	1E-5
随机失活	0.3
全连接网络的维度	128
批处理大小	16

#### 3.3 评价指标

本文使用准确率 (Accuracy) 和宏平均 F1 值 (MF1) 对论辩关系预测模型的性能进行评估. 两个评价指标的公式如下所示:

$$Accuracy = \frac{T}{N} \quad (16)$$

$$F_k = \frac{2P_k R_k}{P_k + R_k} \quad (17)$$

$$MF1 = \frac{1}{K} \sum_{k=1}^K F_k \quad (18)$$

其中,  $T$  表示模型预测正确的样本数量,  $N$  表示输入样本的数量,  $P_k$  代表第  $k$  个类别的精确率,  $R_k$  代表第  $k$  个类别的召回率,  $F_k$  代表第  $k$  个类别的  $F1$  值.

### 3.4 对比方法

为了评估本文提出的模型在 CreateDebate 数据集上的表现, 我们将本文的 CCRnet 模型与目前主流的基线方法和先进方法进行对比.

#### • 基线方法

(1) LSTM-LSTM<sup>[18]</sup>: 基于 Encoder-Decoder 框架的条件模型. 编码器和解码器使用 LSTM 对两个文本分别进行语义表示. 其中, 解码器利用编码器的末状态的输出进行初始化. 模型根据解码器末状态的输出判断论辩关系.

(2) Siamese Network<sup>[20,21]</sup>: 基于 BiLSTM 的孪生神经网络模型. 模型通过对 BiLSTM 编码后的文本语义表示进行拼接, 得到整体语义表示. 模型根据整体语义表示判断交互文本的论辩关系.

#### • 先进方法

(1) ESIM<sup>[34]</sup>: 基于注意力机制的交互推理模型. 模型通过软注意力机制对前提和假设进行局部推理, 并使用 BiLSTM 对它们的局部信息分别进行编码, 得到对应的上下文表示. 模型拼接池化后的上下文表示作为句子关系特征.

(2) Hybrid Network<sup>[22]</sup>: 基于混合注意力机制的论辩关系预测模型. 模型通过自注意力机制和交叉注意力机制提取交互文本各自的关键语义信息和语义关联信息. 模型根据拼接后的关键语义信息和语义关联信息判断交互文本的论辩关系.

(3) BERT<sup>[24]</sup>: 基于双向 Transformer 的预训练语言模型. 其可以应用于句子关系判断、文本分类和阅读理解等下游任务. 在句子关系判断任务中, 模型将最后一层 [CLS] 位置的输出作为特征向量, 用于判断句子关系.

## 4 实验结果与分析

本节首先分析了不同语义表示网络的性能表现. 接着, 将提出的 CCRnet 模型与目前主流的基线方法和先进方法进行对比, 检验模型的有效性. 然后, 设计消融实验评估模型网络结构部件的表现. 最后, 通过可视化相似度矩阵和分析预测错误的样本进一步评估本文的模型.

### 4.1 不同语义表示网络的性能分析

为了分析不同语义表示网络的性能表现, 本文对基于不同语义表示网络的模型性能进行比较. 如表 3 所示, 本文比较了基于 LSTM<sup>[29]</sup>、BiLSTM<sup>[36]</sup>、TextCNN<sup>[37]</sup>和 BERT<sup>[24]</sup>这 4 类语义表示网络的孪生神经网络模型在 CreateDebate 数据集上的性能表现. 其中, 孪生神经网络模型使用拼接的方式融合表述和应答的文本语义表示. LSTM、BiLSTM、TextCNN 均使用 GloVe 词向量<sup>[38]</sup>进行词嵌入表示. LSTM、BiLSTM 均使用末状态的输出作为语义表示向量; TextCNN 使用拼接的池化层输出作为语义表示向量; BERT 使用最后一层 [CLS] 位置的输出作为语义表示向量.

从表 3 中的实验结果可以看出, 基于 BERT 语义表示网络的模型在 CreateDebate 数据集上的整体性能表现最优. 与其他语义表示网络相比, 基于 BERT 语义表示网络的模型在 CreateDebate 数据集上的准确率至少提升了 3.361%,  $MF1$  值至少提升了 3.142%, 说明经过海量语料预训练的 BERT 语言模型具有更强的文本语义表示能力, 能够显著提升模型的预测性能. 在 4 类语义表示网络中, 基于 LSTM 语义表示网络的模型在 CreateDebate 数据集上的整体性能表现最弱, 表明编码上下文信息的语义表示网络的性能优于仅编码前向信息的语义表示网络. 此外, 基于 TextCNN 语义表示网络的模型在 CreateDebate 数据集上的表现优于基于循环神经网络的模型, 造成这一结果可能的原因是在线讨论包含具有复杂语义表达的表述和应答文本. 相比于循环神经网络, TextCNN 使用 3 个不同大小的卷积核进行文本语义表示, 能够捕获到更多维度的局部语义特征.



表 3 基于不同语义表示网络的模型性能 (%)

语义表示网络	<i>Accuracy</i>	<i>MF1</i>
LSTM	59.497	56.462
BiLSTM	60.304	57.186
TextCNN	60.647	58.384
BERT	<b>64.008</b>	<b>61.526</b>

#### 4.2 模型表现

为了验证 CCRnet 模型的有效性,我们将 CCRnet 模型与目前主流的方法进行对比.模型在 CreateDebate 数据集上的实验结果如表 4 所示.

表 4 不同模型在 CreateDebate 数据集上的性能 (%)

模型	<i>Accuracy</i>	<i>MF1</i>
LSTM-LSTM <sup>[18]</sup>	57.042	51.797
Siamese Network <sup>[20,21]</sup>	60.304	57.186
ESIM <sup>[34]</sup>	60.072	56.298
Hybrid Network <sup>[22]</sup>	60.339	57.354
BERT <sup>[24]</sup>	62.475	59.874
CCRnet (Ours)	<b>65.297</b>	<b>63.540</b>

从表 4 中的实验结果可以看到,本文提出的论辩关系预测模型在 CreateDebate 数据集上的准确率和 *MF1* 值均高于其他对比模型,说明本文提出的 CCRnet 模型是有效的.在对比方法中,LSTM-LSTM 模型在 CreateDebate 数据集上的整体性能表现明显弱于其他方法.这是由于该模型仅根据应答文本的条件语义向量判断论辩关系,没有充分利用两个文本的语义信息.Siamese Network、ESIM 和 Hybrid Network 在 CreateDebate 数据集上的性能表现差异不大,但明显弱于 BERT 和本文提出的 CCRnet 模型,说明语义表示网络的编码能力对于整个模型的性能有着重要影响.与 BERT 模型相比,本文提出的模型在 CreateDebate 数据集上的准确率提升了 2.822%,*MF1* 值提升了 3.666%.这是由于 BERT 模型仅使用最后一层 [CLS] 位置的输出判断论辩关系,无法完整地表征两个文本之间的关系特征.与表 3 中基于 BERT 语义表示网络的孪生神经网络模型相比,本文提出的模型在 CreateDebate 数据集上的准确率提升了 1.289%,*MF1* 值提升了 2.014%,表明挖掘交互文本的语境关联信息是有效的,能够进一步提升模型的性能.

#### 4.3 网络结构部件对于模型性能的影响

本文设计消融实验评估模型的不同网络结构部件对于模型性能的影响.具体地,本文单独移除模型中的全连接网络、基于时间分布的全连接网络、表述到应答的注意力、应答到表述的注意力和双向注意力进行实验,并与原有模型进行比较.具体的实验结果如表 5 所示.

表 5 不同网络结构部件对模型性能的影响 (%)

模型	<i>Accuracy</i>	<i>MF1</i>
CCRnet (Ours)	<b>65.297</b>	<b>63.540</b>
移除Dense	64.641	62.553
移除TimeDistributed Dense	64.321	62.542
移除 $P$ to $R$ attention	64.780	63.094
移除 $R$ to $P$ attention	64.472	62.318
移除Co-attention	64.182	62.014

从表 5 中的实验结果可以看出,与原有模型相比,移除任何一个网络结构部件都会降低模型的性能表现,验证了这些网络结构部件的有效性.与原有模型相比,移除 TimeDistributed Dense 层使模型在 CreateDebate 数据集上

的准确率降低了 0.976%, *MF1* 值降低了 0.998%, 说明对每一个带有应答感知的表述词信息和应答中每一个词的语义表示进行降维是必要的. 与原有模型相比, 移除双向注意力机制使模型在 *CreateDebate* 数据集上的准确率降低了 1.115%, *MF1* 值降低了 1.526%, 说明挖掘语境关联信息能够帮助模型更好地判断交互文本的论辩关系. 在双向注意力机制中, 移除应答到表述的注意力比移除表述到应答的注意力对模型整体性能的影响更大, 使模型在 *CreateDebate* 数据集上的准确率降低了 0.825%, *MF1* 值降低了 1.222%, 说明捕捉表述中与应答相关联的最重要的词信息对于判断论辩关系是至关重要的.

#### 4.4 可视化相似度矩阵

为了进一步验证双向注意力机制的有效性, 本文对图 1 中例句 4 所示的表述-应答交互文本的相似度矩阵进行可视化. 具体的可视化结果如图 5 所示. 其中, 横坐标为应答文本, 纵坐标为表述文本.

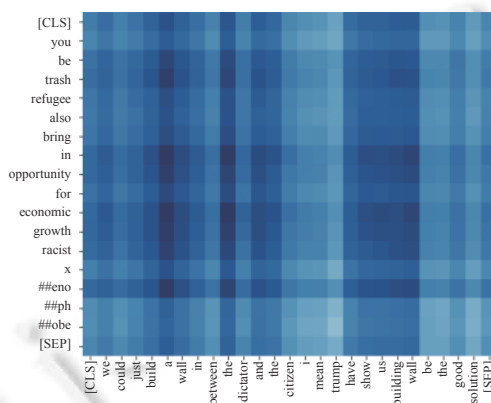


图 5 相似度矩阵可视化

从图 5 可以看出, 双向注意力机制能够捕捉到表述中的关键信息“refugee also bring in opportunity...”和应答中的关键信息“build a wall in between...”是存在关联的. 应答中的“build a wall...”是对表述中“refugee also bring in opportunity...”的反驳, 两者的语境关联信息有助于模型更准确地判断表述和应答的论辩关系. 上述结果表明双向注意力机制能够有效捕捉表述-应答交互文本的语境关联信息并进一步服务于论辩关系预测任务.

#### 4.5 错误分析

为了更全面地评估本文提出的论辩关系预测模型, 我们针对 CCRnet 模型预测错误的交互文本进行分析. 具体地, 本文分别从 *CreateDebate* 测试集的 3 类论辩关系中随机抽取 100 条分类错误的交互文本进行分析. 我们发现隐含的语气表达是模型面临的主要挑战, 影响了模型在 *CreateDebate* 数据集上的表现. 具体包含以下两种情况:

(1) 模型不易判断包含讽刺语气的交互文本之间的论辩关系. 如表 6 的例句 1 所示, 从字面上看, 应答通过“*Yes*”和“*safer*”等单词支持表述中的“no terrorist group will try to come into the U.S.”. 然而, 应答是具有讽刺语气的文本, 其实际含义与字面完全相反. 模型仅根据词汇的语义表达信息很难正确判断两者之间的论辩关系.

表 6 错误示例

编号	示例	标签	预测标签
1	表述: The U.S does through background checks so their is a big possibility that <b>no terrorist group</b> will try to come into the U.S. 应答: <b>Yes</b> a screening that takes up to 24 months. I feel <b>safer</b> already.	反驳	支持
2	表述: Do you even know what racism is? 应答: Do you?	反驳	中立
3	表述: Do you really think companies are off-shoring American jobs to places where they make <b>HIGHER</b> wages? 应答: Not what I said and certainly not why companies are off-shoring.	中立	反驳

(2) 模型不易判断问句形式的文本之间的论辩关系. 如表 6 的例句 2 所示, 应答是以反问的语气对表述进行反驳, 而模型将两者错判为中立关系. 类似地, 如表 6 的例句 3 所示, 应答仅是单纯地对表述中的问题进行回答, 两者为中立关系, 而模型将两者错判为反驳关系. 问句是可以隐晦地表达多种语气的句子类别, 模型从词汇的语义表达中很难直接感知问句的语气, 从而将两者之间的论辩关系判错.

## 5 总结与展望

面向交互文本的论辩关系预测对于自然语言处理领域的观点挖掘、自动摘要等任务的研究具有重要意义. 本文提出了一种挖掘语境关联的双向注意力网络模型 CCRnet, 通过双向注意力机制从表述和应答的交互语境中挖掘两者之间的语义关联信息, 弥补了由于词汇表达的多样性对论辩关系预测带来的不利影响. CCRnet 模型使用共享权重的预训练 BERT 语言模型对表述和应答分别进行文本语义表示, 并通过双向注意力机制建模表述到应答和应答到表述两个方向的注意力, 得到交互文本的语境关联表示. 在此基础上, 模型将表述的文本语义表示和语境关联表示进行融合和池化, 得到最受应答关注的表述特征, 并将其和应答特征进行融合, 建模全局关系特征, 进而输出预测标签. 在 CreateDebate 数据集上的实验结果表明, 与目前主流的基线方法和先进方法相比, 本文提出的 CCRnet 模型的整体性能更优. 此外, 本文通过可视化相似度矩阵进一步验证了双向注意力机制的有效性.

本文提出的模型同样适用于问答匹配和网页检索等任务, 通过挖掘用户查询和候选项之间的关联信息判断两者是否匹配, 进而为用户返回查询结果. 然而, 模型仍然存在一些不足之处, 其无法有效判断含有隐含语气表达的文本之间的论辩关系. 针对这一问题, 未来的工作将考虑利用交互双方的用户特性和交互文本所在的上下文信息作为背景知识, 辅助现有的模型判断两个文本的论辩关系. 此外, 因交互评论文本所属的辩题类型差异较大且缺乏有效信息, 本文没有借助辩题来判断论辩关系. 未来, 我们准备收集具有有效信息的辩题以及相关的评论文本数据, 利用辩题信息作为中间变量来判断交互文本的论辩关系.

## References:

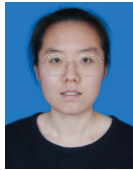
- [1] Stab C, Gurevych I. Identifying argumentative discourse structures in persuasive essays. In: Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014. 46–56. [doi: 10.3115/v1/D14-1006]
- [2] Nguyen H, Litman D. Context-aware argumentative relation mining. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin: Association for Computational Linguistics, 2016. 1127–1137. [doi: 10.18653/v1/P16-1107]
- [3] Stab C, Gurevych I. Parsing argumentation structures in persuasive essays. Computational Linguistics, 2017, 43(3): 619–659. [doi: 10.1162/COLI\_a\_00295]
- [4] Persing I, Ng V. End-to-end argumentation mining in student essays. In: Proc. of the 2016 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego: Association for Computational Linguistics, 2016. 1384–1394. [doi: 10.18653/v1/N16-1164]
- [5] Potash P, Romanov A, Rumshisky A. Here's my point: Joint pointer architecture for argument mining. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. Copenhagen: Association for Computational Linguistics, 2017. 1364–1373. [doi: 10.18653/v1/D17-1143]
- [6] Eger S, Daxenberger J, Gurevych I. Neural end-to-end learning for computational argumentation mining. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver: Association for Computational Linguistics, 2017. 11–22. [doi: 10.18653/v1/P17-1002]
- [7] Liao XW, Chen ZZ, Gui L, Cheng XQ, Chen GL. An argumentation mining method based on multi-task iterative learning. Chinese Journal of Computers, 2019, 42(7): 1524–1538 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2019.01524]
- [8] Liao XW, Ni JC, Wei JJ, Wu YB, Chen GL. Argumentation mining based on multi-task joint learning. Pattern Recognition and Artificial Intelligence, 2019, 32(12): 1072–1079 (in Chinese with English abstract). [doi: 10.16451/j.cnki.issn1003-6059.201912002]
- [9] Chen Y, Cheng XQ, Yang S. Finding high quality threads in Web forums. Ruan Jian Xue Bao/Journal of Software, 2011, 22(8): 1785–1804 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3857.htm> [doi: 10.3724/SP.J.1001.2011.03857]
- [10] Ma HF, Zhang D, Zhao WZ, Shi ZZ. Microblog recommendation method based on hypergraph random walk tag extension. Ruan Jian Xue Bao/Journal of Software, 2019, 30(11): 3397–3412 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5545.htm> [doi: 10.13328/j.cnki.jos.005545]

- [11] Jin ZW, Cao J, Zhang YD, Luo JB. News verification by exploiting conflicting social viewpoints in microblogs. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI Press, 2016. 2972–2978.
- [12] Mukherjee A, Liu B. Discovering user interactions in ideological discussions. In: Proc. of the 51st Annual Meeting of the Association for Computational Linguistics. Sofia: Association for Computational Linguistics, 2013. 671–681.
- [13] Ghosh D, Muresan S, Wacholder N, Aakhus M, Mitsui M. Analyzing argumentative discourse units in online interactions. In: Proc. of the 1st Workshop on Argumentation Mining. Baltimore: Association for Computational Linguistics, 2014. 39–48. [doi: [10.3115/v1/W14-2106](https://doi.org/10.3115/v1/W14-2106)]
- [14] Rosenthal S, McKeown K. I couldn't agree more: The role of conversational structure in agreement and disagreement detection in online discussions. In: Proc. of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue. Prague: Association for Computational Linguistics, 2015. 168–177. [doi: [10.18653/v1/W15-4625](https://doi.org/10.18653/v1/W15-4625)]
- [15] Hassan A, Abu-Jbara A, Radev D. Detecting subgroups in online discussions by modeling positive and negative relations among participants. In: Proc. of the 2012 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Jeju Island: Association for Computational Linguistics, 2012. 59–70.
- [16] Wang L, Cardie C. Improving agreement and disagreement identification in online discussions with a socially-tuned sentiment lexicon. In: Proc. of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Baltimore: Association for Computational Linguistics, 2014. 97–106. [doi: [10.3115/v1/W14-2617](https://doi.org/10.3115/v1/W14-2617)]
- [17] Carstens L, Toni F. Towards relation based argumentation mining. In: Proc. of the 2nd Workshop on Argumentation Mining. Denver: Association for Computational Linguistics, 2015. 29–34. [doi: [10.3115/v1/W15-0504](https://doi.org/10.3115/v1/W15-0504)]
- [18] Bosc T, Cabrio E, Villata S. Tweeties squabbling: Positive and negative results in applying argument mining on social media. In: Proc. of the 6th Biennial Int'l Conf. on Computational Models of Argument. Potsdam: COMMA, 2016. 21–32.
- [19] Chakrabarty T, Hidey C, Muresan S, McKeown K, Hwang A. AMPERSAND: Argument mining for persuasive online discussions. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 2933–2943. [doi: [10.18653/v1/D19-1291](https://doi.org/10.18653/v1/D19-1291)]
- [20] Cocarascu O, Toni F. Identifying attack and support argumentative relations using deep learning. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. Copenhagen: Association for Computational Linguistics, 2017. 1374–1379. [doi: [10.18653/v1/D17-1144](https://doi.org/10.18653/v1/D17-1144)]
- [21] Cocarascu O, Toni F. Combining deep learning and argumentative reasoning for the analysis of social media textual content using small data sets. *Computational Linguistics*, 2018, 44(4): 833–858. [doi: [10.1162/coli\\_a\\_00338](https://doi.org/10.1162/coli_a_00338)]
- [22] Chen D, Du JC, Bing LD, Xu RF. Hybrid neural attention for agreement/disagreement inference in online debates. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018. 665–670. [doi: [10.18653/v1/D18-1069](https://doi.org/10.18653/v1/D18-1069)]
- [23] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [24] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics, 2019. 4171–4186. [doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423)]
- [25] Xi XF, Zhou GD. A survey on deep learning for natural language processing. *Acta Automatica Sinica*, 2016, 42(10): 1445–1465 (in Chinese with English abstract). [doi: [10.16383/j.aas.2016.c150682](https://doi.org/10.16383/j.aas.2016.c150682)]
- [26] Bosc T, Cabrio E, Villata S. DART: A dataset of arguments and their relations on twitter. In: Proc. of the 10th Int'l Conf. on Language Resources and Evaluation. Portorož: European Language Resources Association, 2016. 1258–1263.
- [27] Hewett F, Rane RP, Harlacher N, Stede M. The utility of discourse parsing features for predicting argumentation structure. In: Proc. of the 6th Workshop on Argument Mining. Florence: Association for Computational Linguistics, 2019. 98–103. [doi: [10.18653/v1/W19-4512](https://doi.org/10.18653/v1/W19-4512)]
- [28] Mann WC, Thompson SA. Rhetorical structure theory: Toward a functional theory of text organization. *Text-interdisciplinary Journal for the Study of Discourse*, 1988, 8(3): 243–281. [doi: [10.1515/text.1.1988.8.3.243](https://doi.org/10.1515/text.1.1988.8.3.243)]
- [29] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- [30] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, 1(4): 541–551. [doi: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541)]
- [31] Seo M, Kembhavi A, Farhadi A, Hajishirzi H. Bidirectional attention flow for machine comprehension. arXiv: 1611.01603, 2018.
- [32] Du JC, Xu RF, He YL, Gui L. Stance classification with target-specific neural attention networks. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. Melbourne: AAAI Press, 2017. 3988–3994.

- [33] Zhou YW, Cristea AI, Shi L. Connecting targets to tweets: Semantic attention-based model for target-specific stance detection. In: Proc. of the 18th Int'l Conf. on Web Information Systems Engineering. Puschino: Springer, 2017. 18–32. [doi: [10.1007/978-3-319-68783-4\\_2](https://doi.org/10.1007/978-3-319-68783-4_2)]
- [34] Chen Q, Zhu XD, Ling ZH, Wei S, Jiang H, Inkpen D. Enhanced LSTM for natural language inference. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver: Association for Computational Linguistics, 2017. 1657–1668. [doi: [10.18653/v1/P17-1152](https://doi.org/10.18653/v1/P17-1152)]
- [35] Kingma D, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2017.
- [36] Mikolov T, Karafiát M, Burget L, Černocký J, Khudanpur S. Recurrent neural network based language model. In: Proc. of the 11th Annual Conf. of the Int'l Speech Communication Association. Makuhari: ISCA, 2010. 1045–1048.
- [37] Kim Y. Convolutional neural networks for sentence classification. In: Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014. 1746–1751.
- [38] Pennington J, Socher R, Manning CD. GloVe: Global vectors for word representation. In: Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014. 1532–1543. [doi: [10.3115/v1/D14-1162](https://doi.org/10.3115/v1/D14-1162)]

#### 附中文参考文献:

- [7] 廖祥文, 陈泽泽, 桂林, 程学旗, 陈国龙. 基于多任务迭代学习的论辩挖掘方法. 计算机学报, 2019, 42(7): 1524–1538. [doi: [10.11897/SP.J.1016.2019.01524](https://doi.org/10.11897/SP.J.1016.2019.01524)]
- [8] 廖祥文, 倪继昌, 魏晶晶, 吴运兵, 陈国龙. 基于多任务联合学习的论辩挖掘. 模式识别与人工智能, 2019, 32(12): 1072–1079. [doi: [10.16451/j.cnki.issn1003-6059.201912002](https://doi.org/10.16451/j.cnki.issn1003-6059.201912002)]
- [9] 陈友, 程学旗, 杨森. 面向网络论坛的高质量主题发现. 软件学报, 2011, 22(8): 1785–1804. <http://www.jos.org.cn/1000-9825/3857.htm> [doi: [10.3724/SP.J.1001.2011.03857](https://doi.org/10.3724/SP.J.1001.2011.03857)]
- [10] 马慧芳, 张迪, 赵卫中, 史忠植. 基于超图随机游走标签扩充的微博推荐方法. 软件学报, 2019, 30(11): 3397–3412. <http://www.jos.org.cn/1000-9825/5545.htm> [doi: [10.13328/j.cnki.jos.005545](https://doi.org/10.13328/j.cnki.jos.005545)]
- [25] 奚雪峰, 周国栋. 面向自然语言处理的深度学习研究. 自动化学报, 2016, 42(10): 1445–1465. [doi: [10.16383/j.aas.2016.c150682](https://doi.org/10.16383/j.aas.2016.c150682)]



单华玮(1995—), 女, 博士, 主要研究领域为自然语言处理.



路冬媛(1984—), 女, 博士, 副教授, 博士生导师, CCF 专业会员, 主要研究领域为网络数据挖掘, 自然语言处理.