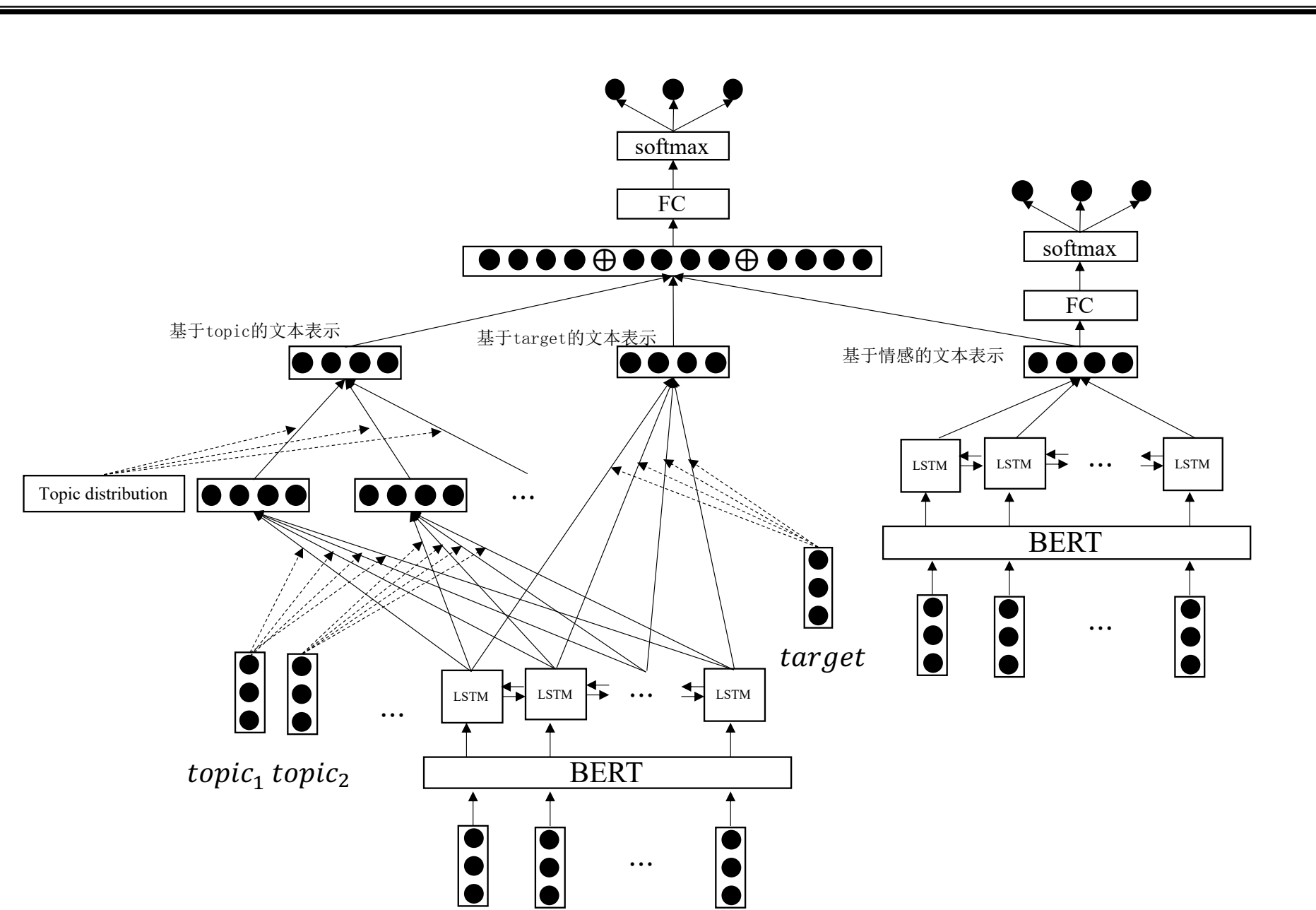


## 论文摘要

• 社交文本立场检测旨在判断给定推文文本针对给定目标所表达的立场倾向性。现有方法主要采用注意力机制来学习面向目标的句子表达。但是, 社交媒体短文本语义表示困难, 且目标缺少有效描述消息。针对上述问题, 本文分别引入BERT预训练模型和LDA模型来增强推文文本的表示和目标的表示, 并利用多注意力机制来建模推文文本与目标主题之间的注意力匹配。此外, 本文进一步引入情感辅助分类器来提取文本中的情感信息。实验表明, 该模型在SemEval-2016立场检测数据集的5个目标上均取得了较好的效果, 微平均F1值提高了约0.75%。

## 模型方法



• 本文提出一种基于情感辅助的多注意力Bi-LSTM立场检测模型。该模型主要分为三部分, 分别为基于主题的文表示, 基于立场目标的文表示和基于情感的文表示。  
• 立场目标信息:  
- 采用注意力机制生成立场目标信息的嵌入表示, 以便关注目标信息中的重要语义信息。

## 模型实验

• 本文实验采用SemEval-2016基于推特的英文立场检测任务数据集。  
• 本文与以下以该数据集进行立场检测的模型进行比较:  
- SVM-ngram: Mohammad等提出的使用word n-grams和character n-grams特征进行立场检测的方法。  
- JOINT: Sun等提出的利用情感信息提升立场检测效果的方法。  
- TAN: Du等提出的基于话题目标增强的注意力机制进行立场检测的方法。  
- AJ-JSS-Lex: Ying等提出的利用情感信息和立场词典, 基于注意力机制的立场检测方法。

模型	AT	CC	FM	HC	LA	MacF <sub>avg</sub>	MicF <sub>avg</sub>
BERT-MA-BiLSTM-SA	74.12	54.68	62.12	70.15	66.50	65.51	73.08
BERT-MA-BiLSTM	73.56	53.62	61.34	70.01	65.57	64.90	72.44
BERT-BiLSTM	72.02	50.06	60.73	69.32	62.50	62.92	71.28
BiLSTM	64.88	43.68	57.93	58.81	60.86	57.23	68.44
SVM-ngram	65.19	42.35	57.46	58.63	66.42	58.01	68.98
JOINT	66.78	50.60	59.35	62.47	61.58	60.16	69.22
TAN	59.33	53.59	55.77	65.38	63.72	59.56	68.79
AJ-JSS-Lex	69.22	59.18	61.49	68.33	68.41	65.33	72.33

• 由实验结果可知, 本文的模型在微平均F1和宏平均F1上分别提高了0.75%和0.18%。  
• 为比较模型每个组件的有效性, 本文对每个组件依次进行实验。其中MA表示使用基于主题的多注意力机制的文表示, SA表示使用情感辅助的文表示。通过消除实验可知每个组件的有效性。

## 论文结论

• 本文的主要工作是引入文本主题信息扩充文本的立场目标, 利用文本的情感信息辅助立场检测任务, 同时利用BERT预训练模型丰富文本的语义表示。  
• 通过LDA主题模型获取同一立场目标下的不同主题, 将文本的不同主题信息与文本信息进行注意力匹配得到基于不同主题的文表示, 同时利用文本所属不同主题的概率分布, 得到基于主题的文表示。通过文本的立场目标与文表示进行注意力匹配, 获得基于立场目标的文表示。同时, 利用预训练好的情感分类器, 得到该文本基于情感的文表示。融合上述三种文表示, 进行立场检测分类, 从而得到分类结果。  
• 通过与目前现有基于立场检测任务的模型在SemEval-2016数据集上的实验结果进行对比, 对本文提出模型的有效性进行了验证。  
• 基于目标的立场检测任务, 需要对立场目标与文本之间的关系进行分析, 仅通过训练数据集无法获取与立场目标相关的常识性信息和背景知识, 在今后的工作中, 将继续考虑扩充立场目标信息的方法, 以便进一步提升立场检测任务的效果。

• 目标主题信息:  
- 使用主题模型获取同一目标下立场文本主题信息, 从而丰富目标信息的内容  
- 使用LDA主题模型获取文本的K个主题信息, 对于每个主题 $T^i$ , 采用该主题下前n个词表示该主题, 即 $T^i = (w_1^i, w_2^i, \dots, w_n^i)$ 。对于主题 $T^i$ 的嵌入表示, 本文利用BERT模型生成该主题下每个词 $w_m^i$ 对应的词嵌入表达 $x_m^i \in R^d$ , 然后根据LDA模型所生成该主题下每个主题词的概率 $p_m^i$ 获得每个词的权重, 得到该主题的嵌入表示, 具体形式如下:

$$\alpha_m^i = \frac{\exp(p_m^i)}{\sum_m \exp(p_m^i)}$$

$$T^i = \sum_m \alpha_m^i x_m^i$$

• 多注意力机制:  
- 引入多个注意力机制, 将立场目标和话题表示与文本中的隐层表达进行匹配, 得到文本中每个隐层表达的相关性权重。对于每个主题 $T^i$ , 注意力机制分配权重的形式如下所示:

$$F(h_j, T^i) = \tanh(h_j W_j^i T^i + b_j^i)$$

$$\beta_j^i = \frac{\exp(F(h_j, T^i))}{\sum_j \exp(F(h_j, T^i))}$$

- 再将双向LSTM隐层输出 $[h_1, h_2, \dots, h_n]$ 与注意力权重 $\beta_j^i$ , 加权得到针对于主题 $T^i$ 的文表示 $S^i$ :

$$S^i = \sum_j \beta_j^i h_j$$

- 基于每个文本, LDA模型会生成其属于每个主题的概率 $q_i$ , 因此本文将基于主题的文表示按照该文本的主题分布概率进一步融合, 具体形式如下:

$$S^{topics} = \sum_i q_i S^i$$

- 最终得到基于主题的文表示 $S^{topics}$ 和基于目标的文表示 $S^{target}$ 。

• 情感辅助分类:  
- 预训练情感辅助分类器, 使文本的情感表达参与文本的立场判断。情感辅助分类器采用Bi-LSTM模型训练, 将BERT模型生成的文本词向量 $[x_1, x_2, \dots, x_n]$ , 通过Bi-LSTM模型及注意力机制得到其文表示 $S^{senti}$ 。

• 模型融合:  
- 将基于目标的文表示、基于主题的文表示和基于情感的文表示的拼接  
- 使用全连接层产生最终的分类结果, 再通过softmax函数输出每个类别的预测概率。