

基于多掩码与提示句向量融合分类的立场检测

王正佳,李 霏,姬东鸿,滕 冲

(武汉大学 国家网络安全学院 空天信息安全与可信计算教育部重点实验室,湖北 武汉 430072)

摘要:立场检测是指分析文本对于某一目标话题表达的立场,立场通常分为支持、反对和其他。近期的工作大多采用 BERT 等方法提取文本和话题的句语义特征,通常采用 BERT 首符号隐藏状态或者句子中每个词隐藏状态取平均作为句向量。该文对句向量的获取进行了改进,采用提示学习模板获取提示句向量,提高句向量的特征提取效果。设计了一种基于多掩码与提示句向量融合分类的立场检测模型(PBMSV),将提示句向量分类与多掩码的模板-答案器结构提示学习分类结合,向句向量引入文本、话题和立场词信息,融合句向量和答案器分类结果,对模型进行联合优化。在 NLPCC 中文立场检测数据集上的实验表明,在五个话题单独训练模型的实验中,该文方法与此前最优方法相比在三个目标上取得领先或持平,取得了 79.3 的总 F1 值,与最优方法接近,并在句向量对比实验中,验证了提示句向量的优势。

关键词:立场检测;深度学习;提示学习;句向量;多掩码

中图分类号:TP391.1

文献标识码:A

文章编号:1673-629X(2023)12-0156-07

doi:10.3969/j.issn.1673-629X.2023.12.022

Stance Detection Based on Multi-mask and Prompt Sentence Vector Fusion Classification

WANG Zheng-jia, LI Fei, JI Dong-hong, TENG Chong

(Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China)

Abstract: Stance detection refers to the analysis of the stance expressed by the text on a target topic, which usually includes support, against and none. Existing works mostly use methods such as BERT to extract sentence feature vectors of the text and topic, and usually, the first token hidden state or the average of the hidden states of each word in the sentence is used as the sentence vector. We improve the acquisition of sentence vectors by using prompt learning templates to obtain prompt sentence vectors and enhance the feature extraction effect of sentence vectors. A stance detection model based on multiple masks and prompt sentence vector fusion classification is designed, which combines prompt sentence vector classification with the template-verbalizer structure of prompt learning classification with multiple masks, introducing text, topic, and stance words information into sentence vectors, fusing sentence vectors and verbalizer classification results, and jointly optimizing the model. Experiments on the NLPCC Chinese stance detection dataset show that in the experiments of training separate models for five topics, the proposed method is superior or comparable to the previous best method in three targets, achieving a total F1 value of 79.3, which is close to the best method. The advantage of prompt sentence vectors is verified in the sentence vector comparison experiment.

Key words: stance detection; deep learning; prompt learning; sentence vector; multi-mask

0 引言

立场检测是指分析确定文本作者对目标(如实体、事件、观点等)的立场,通常将立场分为三类,“支持(favor)”“反对(against)”和“其他(none)”^[1]。该任务属于文本分类任务,给定目标(target)和文本(text),判断文本作者对该目标的立场(stance)类别。

立场检测是舆情分析中的重要手段。随着互联网的快速发展,人群越来越多地在社交媒体平台传播和获取信息,并针对于特定热点发表看法、表明立场。通过立场检测,掌握社交媒体用户对于重要目标的立场,是舆情风险控制的重要环节。比如英国数据公司剑桥分析曾在 2016 年美国大选中借助立场检测手段判断选民

收稿日期:2023-03-12

修回日期:2023-07-13

基金项目:国家自然科学基金(62176187)

作者简介:王正佳(1998-),男,硕士研究生,研究方向为自然语言处理;李 霏(1982-),男,博士,副研究员,研究方向为自然语言处理;通讯作者:姬东鸿(1967-),男,博士,教授,研究方向为自然语言处理、社会计算、深度学习。

对候选人的立场,并精准投放广告影响选民立场。因此,利用立场检测掌握社交媒体上民众对于特定事件的立场,反制敌对势力的舆论引导,关乎网络舆情安全乃至国家安全。

立场检测方法早期主要是基于特征工程的机器学习方法,基于词向量、词频等信息,利用 SVM、词袋模型、随机森林、决策树等方法进行立场分类^[2-3]。随着深度学习方法的兴起,特征提取从人工转为自动学习,基于 LSTM、CNN 以及基于大规模预训练模型的方法取得了相当好的效果^[4-5]。在特征的选取方面,提取并融合语义特征、情感特征、主题特征是目前主流的模式设计思路。通过特征提取网络获取文本或主题的句向量是诸多方法的基础,如近期采用 bert 模型获取文本和主题句向量的方法^[6-7]。通常获取 bert 句向量有两种手段,其一使用 bert 特殊符号 [CLS] 的隐层状态作为句向量,该符号为 bert 自带添加于句首,包含一定的全句信息;其二则将句子中每个词对应的 bert 隐层状态进行融合,如对每个词隐层状态求平均作为句向量。

提示学习是自然语言处理最近兴起的方法,通过在输入中添加若干提示(prompt),提高预训练模型的性能^[8-10]。与过去的微调(tuning)方法不同,思路提示学习方法不再是让模型去适应任务,而是通过构造合适的提示使任务向模型靠拢,保留原有的分类器。提示学习方法发挥预训练模型自身学习的知识,通过提示引入语义引导模型生成结果,或者配合提示进行微调训练。掩码语言模型(Masked Language Model, MLM)是提示学习中常用的预训练模型类型(如 bert 模型)^[11-12],通常基于 MLM 的提示学习提示模板中只有一个掩码位置,然而中文词汇或词组常常超过一个字(词),无法全部填入掩码位置,因此常常只将第一个字(词)填入。比如,“作者的态度是__”,假如态度是“支持”,由 bert 分词器分词后包含两个字“支”和“持”,只有一个掩码位置,则通常只填入第一个字“支”,但仅填入一个字不利于形成通顺语义。

针对前文所述要点,该文提出了基于多掩码与提示句向量融合分类立场检测模型。首先,对句向量的生成方法进行改进,构造提示模板利用 bert 生成句向量;其次,在结构上进行创新,设计了多掩码提示学习分类结构,支持多字填入,并构造合适的模板将提示句向量和多掩码提示学习相结合,提示句向量专注于对文本语义的聚合,通过训练新分类器对立场进行隐式的分类,多掩码提示学习分类利用具体的答案器,显式判断立场,同时答案词能为句向量引入立场词信息,加强句向量与具体的如“支持”等立场词的联系。在 NLPCC 中文立场分类数据集上进行了实验。与相关

基线模型相比,该方法效果明显,取得了 79.3 的总 F1 值,与最优方法接近。

1 相关工作

1.1 立场检测

2016 年 Mohammad 等^[13]从推特上收集文本构建了立场检测英文数据集, SemEval-2016 Task6 立场检测任务使用了该数据集作为任务数据集。Xu 等^[1]构建了基于中文微博文本的立场检测数据集,并发布在 NLPCC-2016 会议的立场检测任务。

早期的立场检测方法基于特征工程利用机器学习方法进行分类。Dian 等^[2]提出基于多文本特征融合的立场检测方法,综合词袋特征、共现特征、词向量等多种特征信息,利用支持向量机(Support Vector Machine, SVM)、随机森林等多种方法进行立场分类。在 Zhang 等^[3]的工作中,提取了情感词和主题词作为特征词,再利用对应的词向量使用 SVM 进行分类。

深度学习兴起后利用深度网络提取特征进行立场检测成为主流。Bai 等^[4]设计了基于双向长短时记忆网络(Bidirectional Long Short Term Memory Network, BiLSTM)与卷积神经网络(Convolutional Neural Network, CNN)进行特征提取,经过注意力池化,再连接特征嵌入进行特征融合并分类的方法。Du 等^[5]提出了特定目标的注意力网络立场检测方法,利用双向循环神经网络(Recurrent Neural Network, RNN)编码文本,训练目标特征嵌入,将文本与目标信息进行内积计算融合特征,训练分类器分类。预训练模型出现后相关方法取得了更好的效果,由 bert 生成的句向量相比此前的 RNN 相关方法优势明显。Wang 等^[6]设计的 Bert-Condition-CNN 模型采用 bert 的 [CLS] 作为句向量,构建话题集与句向量间的关系矩阵,用 CNN 提取特征,对立场进行分类。Geng 等^[7]利用 bert 获取话题和文本表示向量,其中话题句向量由词隐层状态平均得到,并用卷积注意力融合特征,提取文本的词语级与句子级情感特征,将特征拼接进行分类,取得了 NLPCC2016 数据集上目前的最佳结果。

1.2 提示学习

提示学习方法通过合适的提示将下游任务转化为适合模型的形式,通常无需训练新分类器,可以在原分类器上微调,在监督任务以及少样本、零样本任务中均有不错的表现^[8]。在提示形式方面, Petroni 等^[11]提出的 LAMA 模型定义了完形填空形式的提示模板, Li 等^[14]使用了前缀形式的提示;在答案设计方面, Schick 等^[12]使用了人工设计的答案器(verbalizer), Hambarzumyan 等^[15]提出的 WARP 框架则使用了新定义的一组嵌入作为答案器。掩码语言模型(Masked

Language Model, MLM)是语言模型的一种,最具有代表性的是 bert 模型^[16],在预训练过程中语料被随机掩蔽部分位置,再由模型还原。单向生成语言模型(Left-to-Right LM,如 GPT-3 模型^[10])、编解码结构语言模型(Encoder-Decoder LM,如 T5 模型^[17])、前缀语言模型(Prefix LM,如 UniLM1 模型^[18])也广泛应用于提示学习。该文使用了完形填空形式的提示模板以及人工设计的答案器。

2 多掩码与提示句向量融合立场检测

2.1 任务定义

对于立场检测任务,输入是一个数据集 $D = \{(x_1, t_1), (x_2, t_2), \dots, (x_n, t_n)\}$, x 表示一个句子, t 表示特定目标, D 包含 n 个样本;对于每一个样本,输出该文本 x 对于目标 t 的立场标签 y ,其中 $y \in Y = \{\text{favor, against, none}\}$,依次表示“支持”“反对”和“其他”三种类别。例如对于句子“春节放鞭炮污染环境,建议禁止”和目标“春节放鞭炮”,应输出立场 $y = \text{against}$ 。

2.2 模型总体框架

如图 1 所示,多掩码与提示句向量融合立场检测模型由四层组成。输入层输入文本(text)和目标(target);提示层是一个经过设计的提示模板,将文本和目标填入,并包含若干掩码位置[mask];bert 编码层得到模板填入内容后的句子,在每个词的位置进行编码;融合分类层包括三个部分,图 1 左侧的 bert 提示句向量分类,由新训练的分类器分类;右侧通过答案器进行分类,不需要训练新分类器;最后由中间的融合概率分类部分对句向量分类和答案器分类结果进行融合,计算损失联合优化模型。

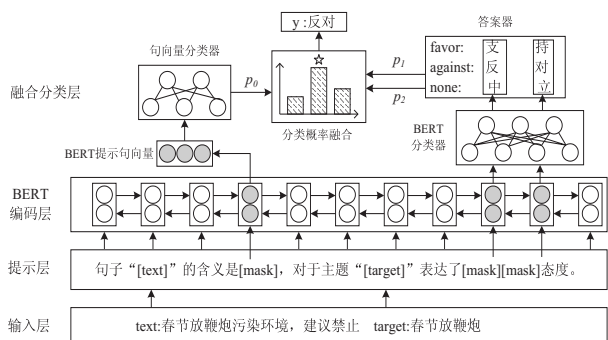


图 1 多掩码与提示句向量融合立场检测模型

2.3 提示层

提示层由一个完形填空形式的提示模板构成,如图 1 所示。模板中的 [text] 和 [target] 位置对应填入数据样本中的文本和目标。

在“句子“[text]”的含义是[mask]”这一句中,[mask]掩码位置用于句向量的生成,提示模板要求模型在 [mask] 处总结句子含义,以此生成能够代表

[text]语义的句向量。

在“对于主题“[target]”表达了[mask][mask]态度”这一句中,两个[mask]掩码位置对应于答案器中的答案词,如“反”和“对”。

由于 bert 模型在其注意力部分允许模型同时看到前后所有位置的信息,因此句向量生成位置虽然在前,也能够捕获主题信息和答案器引入的信息。

2.4 bert 编码层

该文使用预训练模型 bert 作为提示学习框架中的掩码语言模型。对于句子 x ,由 bert 分词器对 x 进行分词,得到 n 个单词,并自动在句子头尾分别添加 [CLS]和[SEP]特殊符,得到 x' 。将分词处理后的句子输入 bert 模型,bert 模型对上下文信息进行综合,在每一个词的位置计算隐层表示,得到全句编码 H :

$$H = \text{BERT}(x') \quad (1)$$

其中, $H = \{H_0, H_1, \dots, H_{n+1}\}$ 。bert 包括多个隐层,对于第 i 位置的词,其对应的隐层可以表示为 $H_i = \{h_0, h_1, \dots, h_j\}$, j 表示有 j 层隐藏状态。

由于本方法只关心 [mask] 位置的特征,如图 1 中包含三个 [mask] 位置,直接用序号 0,1,2 表示,则它们隐藏状态依次为 H_0, H_1, H_2 ,下同。

2.5 融合分类层

2.5.1 bert 提示句向量分类

如图 1,在第一个 [mask] 位置获取 bert 提示句向量。此位置通过提示模板的构造,引导模型概括句子含义。相比于采用 [CLS] 隐藏状态的方法,提示句向量语义指向性更强;而相比于对句子所有词的隐藏表示取平均的做法,bert 提示句向量利用了模型自身的注意力机制,避免了简单平均造成的重要信息权重不足。该 bert 提示句向量记为 s_0 , s_0 由隐藏状态的后 k 层求平均获得,即:

$$s_0 = \frac{1}{k} \sum_{i=j-k+1}^j h_i \quad (2)$$

其中, $h_i \in H_0, k \leq j$ 。该文采用了 $k = 4$,即后四层隐藏状态的平均。

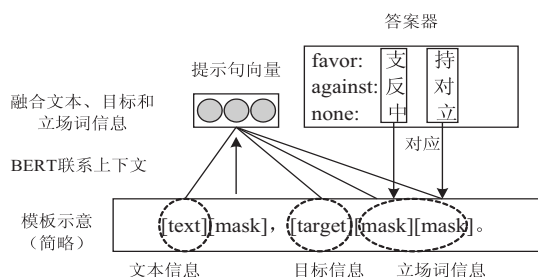


图 2 提示句向量的原理示意图

如图 2 所示,提示句向量由 bert 在图中第一个 [mask] 位置给出,bert 联系上下文综合文本、目标和立场词信息,将特征融合。图 2 中的模板是简略示意,

实际模板是符合语义的完整句子(如图 1),相比于简单拼接,使用提示模板能够降低模型的困惑。

该文采用了一个双层前向神经网络(Feed Forward Neural Networks, FFNN)作为分类器,将 bert 提示句向量输入分类器得到句向量分类概率分布 p_0 ,即:

$$o = \text{relu}(W_0 \cdot s_0 + b_0) \quad (3)$$

$$p_0 = \text{softmax}(W_1 \cdot o + b_1) \quad (4)$$

其中, W_0, b_0 和 W_1, b_1 分别表示第一层与第二层的矩阵与偏置,使用了 relu 激活函数。

2.5.2 多掩码提示学习分类

定义 1: M 为掩码语言模型(MLM), V 表示该模型的词汇表,文本可以表示为 $x = \{w_1, w_2, \dots, w_i\}$, $w \in V$ 。对于输入句子 x ,假设其中存在 m 个被掩蔽的词([mask]),定义 $p_M^k(w|x)$ 表示单词 w 在第 k 个掩码位置由模型 M 给出的概率。

定义 2:提示模板 $T(\cdot)$,将 x 映射到带有若干掩码位置的完形填空句子中。一个模板可以有多个文本填充位置, $T(\cdot)$ 的输入可能是多个句子,如图 1。

定义 3:答案器 $v(\cdot)$,将分类标签 y 映射到具体的词组上,经过分词后词数量可能大于 1,因此定义 $v(y) = \{w_1, w_2, \dots, w_m\}$, m 的大小也对应于模板中掩码的数量,用 $v^k(y)$ 表示 w_k , $w_k \in v(y)$,即 $v(y)$ 中第 k 个单词。以图 1 中后半句为例,由于 bert 按单个汉字分词, $v(y = \text{favor}) = \{\text{支}, \text{持}\}$,长度为 2 即 $m = 2$,对应 $v^1(y = \text{favor}) = w_1 = \text{支}$, $v^2(y = \text{favor}) = w_2 = \text{持}$ 。不同标签对应的答案词分词后长度不一致,则使用填充符补足长度。对于不同类别答案器的词组确保是不同的。

图 3 给出了单独使用的多掩码提示学习分类示意图,与图 1 比较,多掩码提示学习分类与提示句向量分类能够自然地结合。

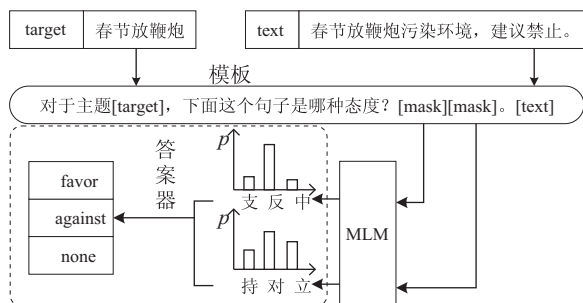


图 3 多掩码提示学习分类示意图

对于掩码位置 k ,分类标签为 y 的概率可以表示为:

$$p(y \in Y | x) = p_M^k(v^k(y) | T(x)) \quad (5)$$

该过程使用模型本身的分类器,无需训练新分类器,记 bert 分类器为 bert_cf。记 $f_v(\cdot)$ 表示根据答案

器 $v(\cdot)$ 所定义的字选取这些字对应位置的值,如图 3 中答案器所示, $f_v(\cdot)$ 从词表概率分布中选出答案词对应的字的概率,以此作为分类依据。向 bert 分类器输入最后一层隐藏状态,同 2.4 节,将图 1 中后两个掩码位置概率分布分别记为 p_1, p_2 ,可得:

$$p_1 = \text{softmax}(f_v(\text{bert_cf}(H_1[-1]))) \quad (6)$$

$$p_2 = \text{softmax}(f_v(\text{bert_cf}(H_2[-1]))) \quad (7)$$

显然不同的 [mask] 位置都可以独立完成分类。

2.5.3 融合概率分类

从句向量分类器和多掩码提示学习两个部分得到了若干概率 p_0, p_1, \dots, p_i ,利用可学习参数 a_0, a_1, \dots, a_i ,将多个概率融合,并使用 softmax 进行归一化,得到最终的概率分布 p 和分类结果 \hat{y} ,即:

$$p = \text{softmax}(\sum_{j=0}^i a_j p_j) \quad (8)$$

$$\hat{y} = \text{argmax}(p) \quad (9)$$

训练时采用交叉熵函数对模型参数 φ 进行优化:

$$L = - \sum_{y \in Y} p(y) \log p(\hat{y}) \quad (10)$$

提示句向量分类器与 bert 指定不同的学习率进行训练,通常提示学习中微调 bert 学习率设置较小。

3 实验设置

3.1 实验数据

采用了立场检测中文数据集 NLPCC2016^[1]对文中方法进行实验测试,包括了 5 个目标上的总计 4 000 条已标注立场类别的中文微博文本数据,并划分了其中 3 000 条作为训练集,1 000 条作为测试集。数据集统计信息如表 1,括号内为目标的简写,与后文实验表格对应。

表 1 NLPCC2016 数据集统计信息

目标	FAVOR	AGAGNST	NONE
iPhone SE (SE)	320	313	167
春节放鞭炮(春节)	338	344	118
俄罗斯在叙利亚的反恐行动(反恐)	344	336	120
开放二胎(二胎)	359	295	146
深圳禁摩限电(禁摩)	223	410	167

3.2 实现细节

文中方法的实现使用了 Pytorch^[19]和 Huggingface Transformers 工具^[20]。掩码语言模型选择了 bert^[16]模型(bert-base-chinese),该模型由 12 层 transformer^[21]编码器组成,词向量维度 768。提示句向量分类器维度为 $768 * 768 * 3$ 。模型训练优化使用了 SGD 优化器。对于句向量分类器学习率设置为 0.01, bert 的参

数学习率设置为 $5e-4$ 。将训练集的 15% 作为验证集选取最佳模型,对不同目标的样本单独训练模型。提示模板、答案器设置如图 1。

3.3 基线模型

机器学习方法:

Dian^[2]:该方法基于词频统计的特征和文本深度特征,使用支持向量机,随机森林和梯度提升决策树对特征进行立场分类。

深度学习方法:

CBL^[22]:CNN-BiLSTM 模型,CNN 提取文本向量的局部特征,BiLSTM 网络提取文本的全局特征,两者结合进行立场检测。

TAN^[5]:该模型通过 BiLSTM 网络与注意力机制提取文本与目标之间的特征信息,实现了基于特定目标的注意力深度分类网络。

BGA^[23]:基于 GCN(卷积图网络)与 BiLSTM, BiLSTM 捕获句子的特征,根据依存句法树建立 GCN,结合注意力得分进行分类。

CTB^[24]:基于 bert 编码文本和话题,将两者拼接,并引入情感分类辅助任务。

BCC^[6]:该模型对微博文本进行主题短语提取来扩充目标话题,利用 Bert 模型获取文本的句向量,构建话题集合与文本句向量间的 Condition 层关系矩阵,用 CNN 提取特征,对立场类别进行分类。

BSECA^[7]:该模型基于 bert 句向量,利用卷积注

意力对文本和话题进行信息提取,融合词语级和句子级情感表示,形成最终的语义表示检测文本立场。该模型额外使用了微博情感数据集训练情感分类器。

3.4 评价指标

按照此前的中文立场检测研究的评价指标惯例,对于特定目标,对支持和反对两个类别计算 F1 值,取二者平均作为该目标上的评价指标:

$$F1 = \frac{F1_{favor} + F1_{against}}{2} \quad (11)$$

总评价指标使用各个目标上 F1 的算术平均值:

$$F1_{total} = \frac{1}{c} \sum F1 \quad (12)$$

其中, c 表示目标类别的数量,近年来的中文立场检测相关研究使用了式 12 衡量总体效果,与 NLPCC 立场检测任务^[1]评价指标设置一致。

4 实验结果与分析

本节表中 PBMSV (Prompt Bert with Multi-mask and Sentence Vector) 表示该文提出的多掩码与提示句向量融合立场检测模型。

4.1 不同方法对比分析

表 2 给出了文中方法与相关对比方法在 NLPCC 中文数据集上对五个目标的数据单独训练模型的实验结果。

表 2 NLPCC 中文数据集五个目标单独训练 (F1)

模型	SE	春节	反恐	二胎	禁摩	Avg.
Dian	61.5	77.6	62.0	84.7	78.2	72.8
CBL	49.4	76.2	50.7	72.2	70.1	63.7
TAN	77.5	59.3	59.2	65.0	72.4	66.7
BGA	63.2	80.5	56.3	84.5	81.4	73.2
CTB	60.6	82.9	59.3	81.9	82.1	73.4
BCC	63.1	80.3	63.6	84.9	80.0	74.4
BSECA	71.0	86.1	68.9	86.2	85.2	79.5
PBMSV	73.4	83.8	68.9	86.9	83.6	79.3

从表 2 可以看出,在该中文数据集上五个目标分开单独训练的实验中,PBMSV 模型在五个目标中有两个目标取得了最好的 F1 值,并取得了接近最优的总 F1 值。相较于采用机器学习方法以及 BiLSTM 获取句子信息的 Dian, TAN, CBL, BGA 方法,采用 bert 生成句向量特征的方法总 F1 提高较多,表现出 bert 预训练模型在特征提取方面的显著优势。采用 bert 进行编码的方法中, BCC 对数据进行充分清洗,采用 condition 层和 CNN 进一步融合特征,相比此前的方法优势明显。

BSECA 设计了更为复杂的特征提取网络,并且利用了额外的微博情感数据集训练单独模型引入情感信息,大幅提高了此前的结果。PBMSV 方法利用了新的提示句向量代替此前的 [CLS] 或平均句向量方法,利用多掩码提示学习分类和提示句向量相结合的方法,提高句向量生成质量,思路上也融合了目标和文本语义,又利用句向量和人工答案器提示学习相结合将“支持”“反对”等具体的立场词信息引入,实验结果在两个目标上超过了 BSECA 方法,一个目标 F1 相同,取得 79.3 的平均 F1 值。考虑 BSECA 方法使用了额

外的微博情感数据集训练情感特征提取模型,文中方法效果上与之接近。

4.2 消融对比实验与分析

为验证 PBMSV 模型中各个模块的有效性,设计了六个消融对比实验,比较各个模块的作用。

PBMSV:主要包括提示句向量 (Prompt Sentence Vector, PSV)、多掩码提示学习 (Prompt with Multi-mask) 和融合分类 (fusion classification, fc) 三个部分。

-fc:去掉融合分类,将句向量分类损失和答案器两掩码位置损失相加进行优化,分类由句向量分类器完成。

-PSV:去掉提示句向量,重新设计提示模板,掩码位置损失相加进行优化,分类由答案器首位置完成。

-PM:去掉多掩码的答案器提示学习分类,保持提示模板不变,由句向量分类器完成分类。

PBSV:由图 1 中的模板前半句生成提示句向量,再将目标输入 bert 模型取 [CLS] 隐藏状态获得目标句向量,将二者拼接输入分类器。

BERT [CLS]:分别将文本和目标输入 bert,它们各自的 [CLS] 隐藏状态作为文本和目标的句向量,将二者拼接输入分类器。

BERT avg.:分别将文本和目标输入 bert,取它们各自每个词的最后一层隐藏状态平均值作为文本和目标的句向量,将二者拼接输入分类器。

表 3 NLPCC 中文数据集上的消融实验 (Avg. F1)

模型	Avg.
PBMSV	79.3
-fc	78.1
-PSV	77.1
-PM	74.8
PBSV	74.3
BERT [CLS]	73.1
BERT avg.	72.4

从表 3 可以看出,两种 BERT 基础句向量实验与 PBSV 比较,结构基本一致,区别在于 PBSV 使用了提示句向量,PBSV 的 F1 值略高表明了提示句向量在特征提取方面有一定的提升效果。**-PM** 和 **PBSV** 比较,两者差距不大,表明了使用模板引入目标信息与直接拼接目标句向量效果接近,而使用模板降低了分类器维度,参数量更少。**-PSV** 只使用多掩码提示学习方法,取得了 77.1 的 F1 值,效果出色,相比于 4.1 节中的多数方法取得领先。**PBMSV** 比 **-fc** 效果更好,验证了融合多个分类结果优化损失的有效性。相比于 **BERT [CLS]** 和 **BERT avg.**,**PBMSV** 分别取得了 8.4% 和 9.5% 的提高。

5 结束语

该文提出了一种基于多掩码与提示句向量融合分类的立场检测模型 (PBMSV)。该模型中对句向量的生成进行了改进,采用提示学习思想,构造语义通顺的模板融合文本、目标和立场词信息,帮助 bert 生成更好的句向量进行分类;采用模板-答案器结构的多掩码提示学习分类方法,使用 bert 自身分类器获取词表概率分布,由答案器分类,与提示句向量配合,在结构上进行创新改进;融合句向量分类结果和答案器分类结果,联合优化模型,完成立场检测分类。在 NLPCC2016 中文立场检测数据集上的相关实验证明了 PBMSV 模型的有效性,并验证了提示句向量的效果。

现有研究大多采用 bert 生成特征,再进一步通过设计网络结构对相关特征进行融合和分类。而提示学习方法着重于发挥预训练模型知识,利用提示帮助模型完成下游任务。除监督方法之外,提示学习方法在少样本场景下也表现出色。在未来的工作中,将进一步地研究提示学习在立场检测中的应用,还将探索少样本场景、迁移场景的立场检测。

参考文献:

- [1] XU R, ZHOU Y, WU D, et al. Overview of nlpcc shared task 4: stance detection in chinese microblogs [C] // Natural language understanding and intelligent applications: 5th CCF conference on natural language processing and chinese computing, NLPCC 2016, and 24th international conference on computer processing of oriental languages, ICCPOL 2016. Kunming: Springer International Publishing, 2016: 907-916.
- [2] 莫雨洁, 金 琴, 吴慧敏. 基于多文本特征融合的中文微博的立场检测 [J]. 计算机工程与应用, 2017, 53 (21): 77-84.
- [3] 郑海洋, 高俊波, 邱 杰, 等. 基于词向量技术与主题词特征的微博立场检测 [J]. 计算机系统应用, 2018, 27 (9): 118-123.
- [4] 白 静, 李 霏, 姬东鸿. 基于注意力的 BiLSTM-CNN 中文微博立场检测模型 [J]. 计算机应用与软件, 2018, 35 (3): 266-274.
- [5] DU J, XU R, HE Y, et al. Stance classification with target-specific neural attention networks [C] // Proceedings of the 26th international joint conference on artificial intelligence. Melbourne: AAAI, 2017: 3988-3994.
- [6] 王安君, 黄凯凯, 陆黎明. 基于 Bert-Condition-CNN 的中文微博立场检测 [J]. 计算机系统应用, 2019, 28 (11): 45-53.
- [7] 耿源羚, 张绍武, 张益嘉, 等. 基于卷积注意力的情感增强微博立场检测 [J]. 山西大学学报: 自然科学版, 2022, 45 (2): 302-312.
- [8] LIU P, YUAN W, Fu J, et al. Pre-train, prompt, and predict:

- a systematic survey of prompting methods in natural language processing[J]. *ACM Computing Surveys*, 2023, 55(9): 1–35.
- [9] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners [J]. *OpenAI Blog*, 2019, 1(8): 9.
- [10] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners [J]. *Advances in Neural Information Processing Systems*, 2020, 33: 1877–1901.
- [11] PETRONI F, ROCKTÄSCHEL T, RIEDEL S, et al. Language models as knowledge bases? [C]//*Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP – IJCNLP)*. Hong Kong: Association for Computational Linguistics, 2019: 2463–2473.
- [12] SCHICK T, SCHÜTZE H. Exploiting cloze-questions for few-shot text classification and natural language inference [C]//*Proceedings of the 16th conference of the european chapter of the association for computational linguistics; main volume*. [s. l.]: Association for Computational Linguistics, 2021: 255–269.
- [13] MOHAMMAD S, KIRITCHENKO S, SOBHANI P, et al. A dataset for detecting stance in tweets [C]//*Proceedings of the tenth international conference on language resources and evaluation (LREC'16)*. Portorož: European Language Resources Association (ELRA), 2016: 3945–3952.
- [14] LI X L, LIANG P. Prefix-tuning: optimizing continuous prompts for generation [C]//*Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: long papers)*. [s. l.]: Association for Computational Linguistics, 2021: 4582–4597.
- [15] HAMBARDZUMYAN K, KHACHATRIAN H, MAY J. WARP: word-level adversarial reprogramming [C]//*Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: long papers)*. [s. l.]: Association for Computational Linguistics, 2021: 4921–4933.
- [16] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//*Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*. Minneapolis: Association for Computational Linguistics, 2019: 4171–4186.
- [17] RAFFEL C, SHAZEER N, ROBERTS A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer [J]. *The Journal of Machine Learning Research*, 2020, 21(1): 5485–5551.
- [18] DONG L, YANG N, WANG W, et al. Unified language model pre-training for natural language understanding and generation [C]//*Proceedings of the 33rd international conference on neural information processing systems*. Vancouver: Curran Associates Inc., 2019: 13063–13075.
- [19] PASZKE A, GROSS S, MASSA F, et al. PyTorch: an imperative style, high-performance deep learning library [C]//*Proceedings of the 33rd international conference on neural information processing systems*. Vancouver: Curran Associates Inc., 2019: 8026–8037.
- [20] WOLF T, DEBUT L, SANH V, et al. Transformers: state-of-the-art natural language processing [C]//*Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*. [s. l.]: Association for Computational Linguistics, 2020: 38–45.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//*Proceedings of the 31st international conference on neural information processing systems*. Long Beach: Curran Associates Inc., 2017: 6000–6010.
- [22] 张翠肖, 郝杰辉, 刘星宇, 等. 基于 CNN-BiLSTM 的中文微博立场分析研究 [J]. *计算机技术与发展*, 2020, 30(7): 154–159.
- [23] 杨顺成, 李彦, 赵其峰. 基于 GCN 和 Bi-LSTM 的微博立场检测方法 [J]. *重庆理工大学学报: 自然科学*, 2020, 34(6): 167–173.
- [24] 岳天驰. 面向微博文本的情绪与立场检测方法研究 [D]. 大连: 大连理工大学, 2020.