

[文章编号]1000-1832(2023)01-0065-07

[DOI]10.16163/j.cnki.dslkxb20210905001

融合 BERT 词嵌入和 BiLSTM 的 微博谣言持续检测模型

何晓霞,古兰拜尔·吐尔洪,买日旦·吾守尔,王松

(新疆大学信息科学与工程学院,新疆 乌鲁木齐 830046)

[摘 要] 针对微博谣言带标签数据不足,且当下的谣言检测模型无法持续学习应对不断变化的微博网络语言等问题,本文提出 BERT-BiLSTM-LML 微博谣言持续检测模型.首先,使用 BERT(Bidirectional Encoder Representations from Transformers)预训练模型提取两个任务输入文本数据的词向量;其次,使用双向长短时记忆(Bi-directional Long Short-Term Memory, BiLSTM)网络充分提取文本的上下文特征;最后,基于 BiLSTM 深层特征使用终身监督学习算法 ELLA(Efficient Lifelong Learning Algorithm)对两个任务的特征数据进行建模,以实现持续检测.实验结果表明:BERT 词向量有效优化了模型性能,比基于 Word2vec 词向量的 Word2vec-BiLSTM-LML 模型在准确率和 F1 值都提升了 5.5%.相较于独立学习,在持续学习争议检测任务后,模型的谣言检测准确率提升了 1.7%,F1 值提升了 1.8%.同时,在持续学习过程中,随着知识的积累,谣言检测准确率持续提升.最终在公开的微博数据集上,BERT-BiLSTM-LML 模型谣言检测准确率为 93.2%,F1 值为 93.1%,优于其他基线模型.

[关键词] 谣言检测;争议检测;终身机器学习;微博

[中图分类号] TP 391

[文献标志码] A

0 引言

随着人类进入互联网时代,越来越多的人开始使用在线社交媒体.在这些平台中,以复杂无序的社交网络为媒介进行信息传播,具有传播范围广和影响力大等特点.在这些传播的信息中不可避免会夹杂着谣言.当谣言传播到一定范围后会给人类的日常生活带来巨大的消极影响,因此,如何有效进行谣言检测(Rumor Detection, RD)具有十分重要的意义.

目前,微博谣言检测领域带标签数据匮乏,研究者通常使用无监督学习^[1]和多任务学习(Multi task learning, MTL)^[2]来解决该问题.但它们存在如下缺陷:(1)无监督学习方法稳定性

差;(2)面对微博平台上不断更新的网络语言,MTL 难以持续学习.在对多个任务训练完后,也难以更新和积累知识帮助学习新的谣言检测任务.针对上述问题,本文构建 BERT-BiLSTM-LML 微博谣言持续检测模型.该模型具体如下:(1)谣言与争议性存在较强关联,故使用争议检测(Controversy Detection, CD)任务辅助谣言检测任务学习^[3],以解决微博谣言带标签数据匮乏的问题,并提升模型的谣言检测效果;(2)使用 BERT 预训练模型获取谣言检测与争议检测输入数据的词向量,并输入到 BiLSTM 网络中进行深层特征提取,以获得更高质量的特征向量;(3)基于 BiLSTM 模型提取的深层特征,使用 LML 领域中的 ELLA 算法对谣言检测与争议检测的特

[收稿日期] 2021-09-05

[基金项目] 新疆维吾尔自治区自然科学基金资助项目(2018D01C075);新疆维吾尔自治区教育厅高校科研青年基金资助项目(61021800032,61021211418);新疆维吾尔自治区高层次人才项目(100400016,042419006).

[作者简介] 何晓霞(1997—),女,硕士研究生,主要从事机器学习、舆情分析研究;通信作者:古兰拜尔·吐尔洪(1985—),女,副教授,主要从事机器学习、网络安全研究.

征数据进行建模,以对这两个相关任务进行持续学习,并在学习过程中,通过不断积累知识实现对微博谣言检测效果的持续改进。

1 相关工作

LML 架构中一般包含知识库 (Knowledge Base, KB) 和基于知识的学习器 (Knowledge-based learner, KBL) 两个组件,如图 1 所示。其中 KB 负责知识的存储,而 KBL 负责学习任务,并进行知识的获取与利用。LML 是一个持续的学习过程,它已学习了 $n-1$ 个任务: $\{T_i\}_{i=1}^{n-1}$, 被称之为旧任务。每个任务都有自己对应的数据集: $\{D_i\}_{i=1}^{n-1}$ 。这些任务可以是相同的领域,也可以是不同的领域。当面对第 n 个任务 T_n 时(被称作当前任务或者新任务),此时的数据集为 D_n , LML 将 KB 中的历史知识进行前向迁移来帮助 KBL 学习当前任务 T_n 。 T_n 学习完成后,KB 将进行知识的存储和更新。LML 的目标是优化新任务或者全部任务的学习效果^[4]。

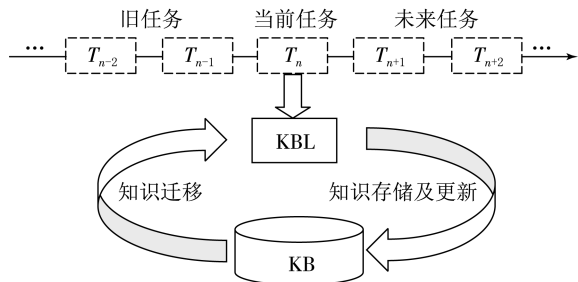


图1 LML学习过程

从上述学习过程看出,LML 和以下两种机器学习方法相关。

(1) 迁移学习。迁移学习是把源领域(任务)的知识迁移到目标领域(任务),使得目标任务能够取得更好学习效果的一种机器学习方法^[5]。它一般用于解决目标任务带标签数据量小的问题。迁移学习只能进行一次知识迁移,且只能从源任务到目标任务。相比之下,LML 进行持续学习,只要新的任务到达,便发生一次知识迁移帮助新任务学习,且在某些 LML 算法上能发生双向的知识迁移,即能同时提升新任务和旧任务的学习效果,例如 ELLA 算法。

(2) 多任务学习。多任务学习同时训练一些相关任务,学习到任务之间的一些共享表示,并进一步地挖掘特定域信息来促进某个任务学习^[6]。多任务学习与 LML 都能使用共享信息帮助任务学习,但不同之处在于多任务学习难以持续学习,也难以进行知识的积累。虽然多任务学

习在新任务到达时通过重新训练能优化所有的任务,但当任务数量足够多时,重新训练将耗费大量的时间和资源,而 LML 能在历史知识的帮助下持续学习多个新任务。

面对微博上不断变化的网络语言,模型的学习应该是持续的。虽然多任务学习能对谣言检测和争议检测这两个相关任务进行联合学习,却难以持续学习新的任务。相比之下,LML 能对两个任务进行持续学习,在学习过程中,随着 KB 中知识的积累和更新,模型持续优化谣言检测效果,所以 LML 更适合应用在微博谣言检测领域。

2 本文模型

本文 BERT-BiLSTM-LML 模型的整体运作流程:给定争议检测及谣言检测这两个任务的微博源帖文本内容,模型利用 BERT 预训练模型,获得包含上下文语义信息的动态词向量,再将该词向量输入到 BiLSTM 神经网络中进行特征提取。将最后时刻的隐含层向量 h_n 作为最终的文本特征向量送入到 ELLA 分类器中,从而达到持续检测微博谣言的目的。BERT-BiLSTM-LML 模型的结构图如图 2 所示。

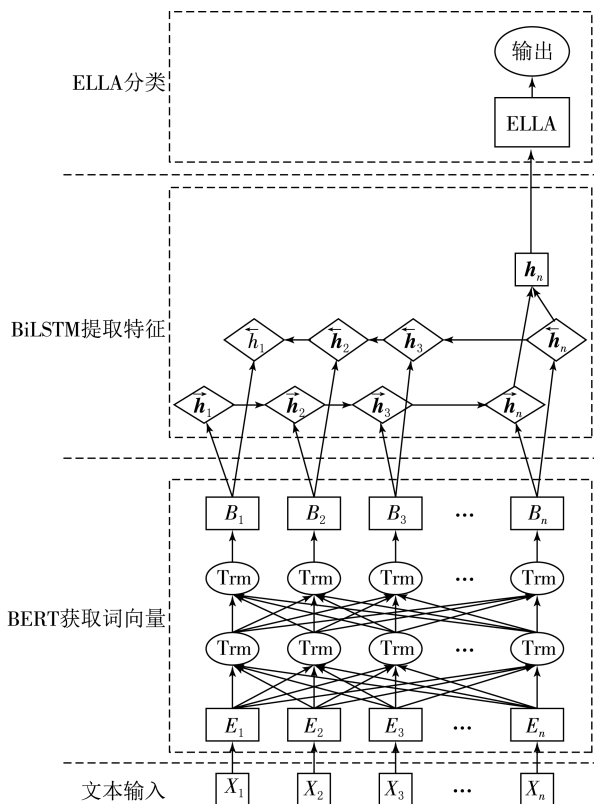


图2 本文模型架构图

2.1 BERT 词嵌入

BERT^[7] 预训练模型得到的词向量可输送到其他模型中进行训练,它的词向量是词嵌入张

量、语句分块张量和位置编码张量做加和的结果. BERT 基于双向的 Transformer 编码器,因此拥有强大的语言表征能力,其结构如图3所示.

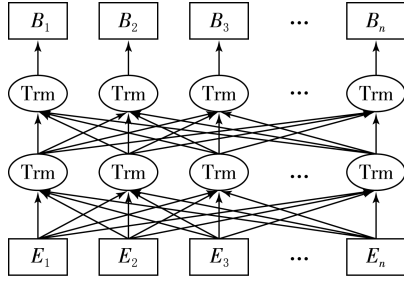


图3 BERT 结构图

2.2 BiLSTM 特征提取

RNN 模型当序列过长时易产生梯度弥散问题,为了克服该缺点,LSTM 被提出^[8]. 它引入了门控机制,通过其特殊的门结构使得模型可以保留文本中的重要特征,而剔除相对无用的信息. 为了充分提取文本的上下文特征,使用 BiLSTM 神经网络模型进行特征提取,它的结构如图4所示. 向量化后的 BERT 词向量为 $\mathbf{B} = \{B_1, B_2, B_3, \dots, B_n\}$, 它分别输入到前向传播层和后向传播层. 前向传播层输出的向量集合为 $\{\vec{h}_1, \vec{h}_2, \vec{h}_3, \dots, \vec{h}_n\}$, 后向传播层输出的向量集合为 $\{\overleftarrow{h}_1, \overleftarrow{h}_2, \overleftarrow{h}_3, \dots, \overleftarrow{h}_n\}$. 其中 \vec{h}_n 为最后时刻前向隐含层输出的特征向量, \overleftarrow{h}_n 为最后时刻后向隐含

层输出的特征向量,将这两个向量进行拼接形成 \mathbf{h}_n 作为最终的文本特征向量.

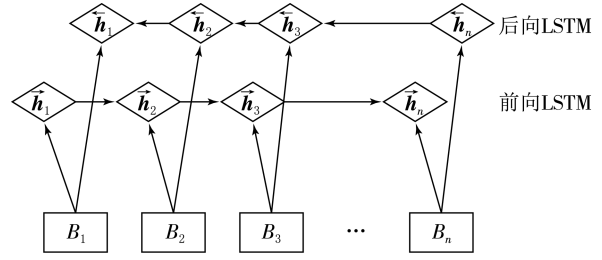


图4 BiLSTM 结构图

2.3 ELLA 持续检测

在使用 BiLSTM 提取争议检测任务和谣言检测这两个任务数据的文本特征后,选取基于 LML 的监督学习算法 ELLA^[9]对两个任务的特征数据进行建模. ELLA 中的 KBL 学习争议检测任务,并产生任务独立知识 (Task-independent Knowledge, TIK) 和共享知识 (Shared Knowledge, SK) 存储在 KB 中,其中的 SK 通过前向知识迁移机制帮助谣言检测任务学习. KBL 学完谣言检测任务后,更新 SK 用于帮助未来任务 (争议检测或谣言检测),并将当前任务产生的 TIK 存储到 KB 中. 随着对这两个任务的持续学习,SK 逐渐更新以不断提升模型的谣言检测效果. 上述持续学习过程如图5所示.

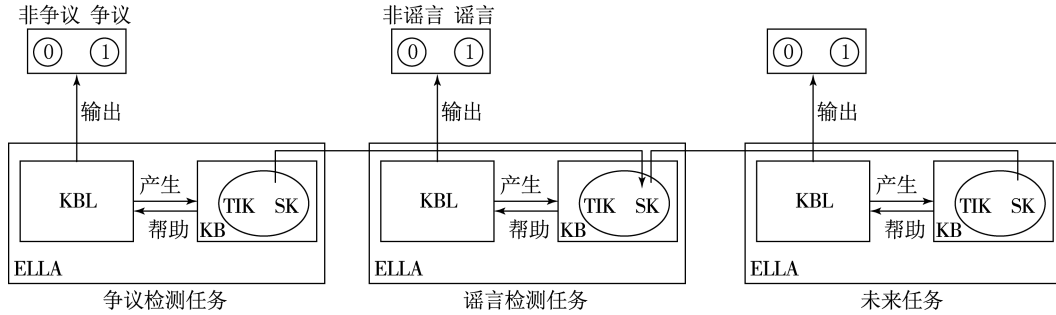


图5 持续学习过程图

ELLA 算法包括 4 个输入参数, k, d, μ 和 λ 分别代表隐含层维数、特征维数、 L_1 范数和 L_2 范数. 设定参数 k 的取值范围是 $k \in [1, \min(10, T/4)]$ (T 为任务总数), λ 设定为 $\{e^{-5}; e^{-2}; e^1; e^4\}$ 中的某个值^[9].

ELLA 算法采用参数化的方法来进行终身学习. 其预测函数定义为

$$f^{(t)}(x) = f(x; \theta^{(t)}) = X_i^t \theta^{(t)}. \quad (1)$$

每个任务都对应一个参数向量 $\theta^{(t)}$ ($\theta^{(t)} \in \mathbb{R}^d$), 也就是任务 t 的预测函数. 它是稀疏共享矩阵 L ($L \in \mathbb{R}^{d \times k}$) 和 $S^{(t)}$ ($S^{(t)} \in \mathbb{R}^k$) 的线性组合, 即

$$\theta^{(t)} = LS^{(t)}. \quad (2)$$

ELLA 算法通过降低在全部任务上的预测损失值进行优化. 目标函数定义为

$$g(L) = \frac{1}{T} \sum_{t=1}^T \min_{s^{(t)}} \left\{ \frac{1}{n_t} \|\theta^{(t)} - LS^{(t)}\|_{D^{(t)}}^2 + \mu \|S^{(t)}\|_1 \right\} + \lambda \|L\|_2^2. \quad (3)$$

其中 $\theta^{(t)} = \min_{\theta} \frac{1}{n_t} \sum_{i=1}^{n_t} \mathcal{L}(f(x_i^{(t)}; \theta), y_i^{(t)})$ 为任务 t 的最优参数模型. $D^{(t)}$ 为损失函数的 Hessian 矩阵, 用于评估 $\theta^{(t)}$.

给定两个任务的特征数据样本 $\{(X_i^t, y_i^t)\}$ ($i=1, 2, \dots, n_t$), n_t 为任务 t 对应的样本数. 其中

$X_i' = (X_{1i}, X_{2i}, \dots, X_{di})$, y_i' 为此任务中该条样本的标签. 在特征基础上, ELLA 对多个任务进行持续学习的详细步骤如下:

(1) ELLA 算法中的初始参数向量 $\theta^{(t)}$ 由基础学习器线性回归或者逻辑回归方式进行计算, 即图 5 中的 KBL. 争议检测和谣言检测实际上是一个分类问题, 故使用逻辑回归初始化任务 t 的参数向量 $\theta^{(t)}$ 和 $D^{(t)}$. 其计算方式为:

$$\sigma_i^{(t)} = \frac{1}{1 + e^{-\theta^{(t)\top} x_i^{(t)}}}; \quad (4)$$

$$D^{(t)} = \frac{1}{2n_t} \sum_{i=1}^{n_t} \sigma_i^{(t)} (1 - \sigma_i^{(t)}) x_i^{(t)} x_i^{(t)\top}. \quad (5)$$

(2) 使用当前的 $\theta^{(t)}$ 和 $D^{(t)}$ 根据(6)式更新模型参数向量 $S^{(t)}$, 其中

$$S^{(t)} \leftarrow \arg \min_{S^{(t)}} (\mu \| S^{(t)} \|_1 + \| \theta^{(t)} - LS^{(t)} \|_{D^{(t)}}^2). \quad (6)$$

在(6)式中 L 的初始化值为 $d * k$ 的零矩阵.

(3) 使用上一步的 $S^{(t)}$ 更新矩阵 L_{new} , 这一步是为了解决收敛优化问题. 新矩阵 L 的计算方法为

$$L_{\text{new}} \leftarrow \arg \min_L \left(\lambda \| L \|_2^2 + \frac{1}{T} \sum_{t=1}^T \| \theta^{(t)} - LS^{(t)} \|_{D^{(t)}}^2 \right). \quad (7)$$

(4) 根据(2)式更新任务 t 的参数向量 $\theta^{(t)}$, 再根据(1)式得到任务 t 中每条微博的标签.

(5) 重复上述步骤, 直至没有新的任务数据. 以上的稀疏共享矩阵 L 、每个任务对应的 $S^{(t)}$ 向量及训练数据作为知识存储在 KB 中. 随着任务的增加, 更新 $S^{(t)}$ 将耗费越来越大的时空代价, 所以计算每个任务的 $S^{(t)}$ 后, 对其他任务训练时不更新这些 $S^{(t)}$. 全部任务的优化则是通过稀疏共享矩阵 L 的更新来完成. 随着 L 的更新, ELLA 算法具备双向的知识迁移能力, 根据目标函数 $g(L)$, 全部任务的预测损失值将逐渐降低, 继而全部任务的分类准确率逐渐上升^[9]. 由于这一特性, 所以模型在理论上也能提升争议检测任务的学习效果.

图 5 中 KB 存储的 TIK 即为上述的 $S^{(t)}$ 向量及任务的训练数据, SK 则指的是稀疏共享矩阵 L .

3 实验与分析

3.1 实验数据

文献[10]给出争议检测数据集. 该数据集包含 49 个被广泛讨论的话题, 这些时间发生在 2017—2019 年. 它包含 2 个子数据集 intra-topic

和 inter-topic. 本文模型具有迁移学习能力, 不需要太多的带标签争议微博数据, 所以本文选取其中的 intra-topic 子数据集用于本文的争议检测任务. intra-topic 子数据集总共包含 5 676 条微博, 其中包含 2 495 条争议性微博, 3 181 条非争议性微博. 每条微博数据包含的信息有微博 id、微博话题、微博源帖内容、表情信息和标签.

文献[11]报道了微博谣言数据集, 数据采集自新浪微博社区管理中心和新浪微博 API 接口, 总共有 4 664 条微博事件, 其中 2 313 条谣言微博, 2 351 条非谣言微博. 每条微博事件的信息存储在 json 文件中, 它包含的信息有该条微博的转发数、评论数、源帖内容、评论转发内容、发布用户的信息和标签等等.

本文用到的数据仅为 2 个数据集中的微博源帖内容及标签.

3.2 实验过程

3.2.1 数据预处理

为了提高后续特征提取的质量, 首先对两个数据集进行清洗. 本文使用正则表达式去除微博源帖内容中的 @ 及 @ 的内容、url 链接、标点符号、# 引用的话题内容和表情符号等等. 然后对两个任务数据进行训练集和测试集的划分, 比例为 7 : 3. 同时, 为了降低实验结果的偶然性而采用五折交叉验证, 取其平均值作为最终的实验结果.

3.2.2 模型参数设定

(1) BERT 预训练模型参数. 考虑到微博短文本的特性, BERT 预训练模型最大长度 Max_seq_length 设置为 140, 池化策略采用最大池化. 最终得到 BERT 词向量的特征维度为 768. 为了显示 BERT 词向量能提升实验效果, 本文将与 Word2vec 预训练模型得到的静态词向量做对比. Word2vec 处理得到的词向量维度为 300.

(2) BiLSTM 神经网络参数. BiLSTM 神经网络基于 Tensorflow2. 3. 0, 实验硬件环境为 i7-11800H CPU, 该模型的参数设置如表 1 所示.

表 1 BiLSTM 神经网络参数

Parameters	Value
batch_size	64
隐含层维度	200
学习率	0.001
遗忘率	0.5
epoch	30

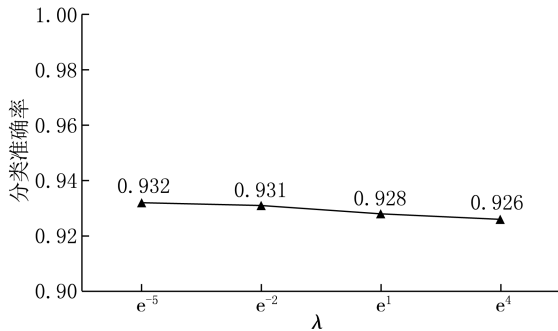
(3) ELLA 持续学习算法参数. 虽然基于 LML 的 ELLA 算法面向的是无穷无尽的任务数

据,但在实验中只能设定为有限的任务数据来验证它的优势.本文需要学习的任务为争议检测和谣言检测,所以 T 设为 2. 由于 BiLSTM 神经网络设置的隐含层维度为 200,所以 ELLA 中的特征维度 d 为 400. ELLA 算法详细的参数设定如表 2 所示.

表 2 ELLA 参数设置

参数	数值
T	2
k	1
μ	0.000 1
d	400
λ	e^{-5}

在表 2 中,当 λ 为 e^{-5} 时, BERT-BiLSTM-LML 模型的谣言检测效果最好,此时的谣言分类准确率达到了 0.932. 不同 λ 下模型的谣言分类效果如图 6 所示.

图 6 λ 参数寻优

3.2.3 评价指标选取

使用 BERT-BiLSTM-LML 模型依次对争议检测任务和谣言检测任务进行训练,训练完后,在谣言检测任务的测试集上进行测试,得到评价指标.评价指标选用准确率(A)和 F1.

3.3 实验对比

3.3.1 Word2vec 与 BERT 词向量对比

为了体现通过 BERT 预训练模型的词向量有利于提升模型的分类效果,将本文模型 BERT-BiLSTM-LML 与基于 Word2vec 词向量的 Word2vec-BiLSTM-LML 模型进行对比,两个模型在谣言检测任务上的实验结果如表 3 所示.

表 3 不同词向量效果对比

Model	A	F1
Word2vec-BiLSTM-LML	0.877	0.876
BERT-BiLSTM-LML	0.932	0.931

从表 3 得知,当基于相同的 BiLSTM 神经网络模型和 ELLA 算法参数, BERT-BiLSTM-LML 模型在谣言检测任务上的分类准确率为

0.932, F1 为 0.931,这两项指标均优于 Word2vec-BiLSTM-LML 模型,所以富含上下文语义信息的 BERT 词向量能提升模型的谣言检测效果.

3.3.2 知识迁移实验结果对比

为了证明学习争议检测任务后产生的知识通过前向知识迁移机制对谣言检测有所帮助,本文将争议检测和谣言检测任务分离,将单独的谣言检测任务送入模型中与其对比,前后的实验结果对比如表 4 所示.其中 RD 指的是模型中只有谣言检测任务时的分类效果,CD+RD 指的是利用争议检测任务产生的知识进行前向知识迁移后,谣言检测任务的分类结果.

表 4 知识迁移效果对比

Task	A	F1
RD	0.915	0.913
CD+RD	0.932	0.931

从表 4 看出:当模型中只有谣言检测任务时谣言检测准确率为 0.915, F1 为 0.913.而发生前向知识迁移后,谣言检测任务上的准确率达到 0.932, F1 达到 0.931,均比未发生知识迁移时有所提升,这说明争议检测任务学习后产生的知识对谣言检测任务帮助性较大.

3.3.3 知识积累对谣言检测的影响

为了验证争议检测知识(KB_CD)和谣言检测知识(KB_RD)在模型学习过程中是如何影响谣言检测效果,本文将用实验结果来探究随着 KB_CD 和 KB_RD 知识的累积,谣言检测任务分类准确率的变化过程,其结果如图 7 所示.其中横坐标代表训练实例的数量.随着训练实例的增加,相关知识进行积累.其中图 7 中的实线是在使用全部的谣言检测训练集时,探索随着 KB_CD 知识的积累谣言检测效果的变化.而虚线是在使用全部的争议检测任务训练集时,探索随着 KB_RD 知识的积累谣言检测效果的变化.

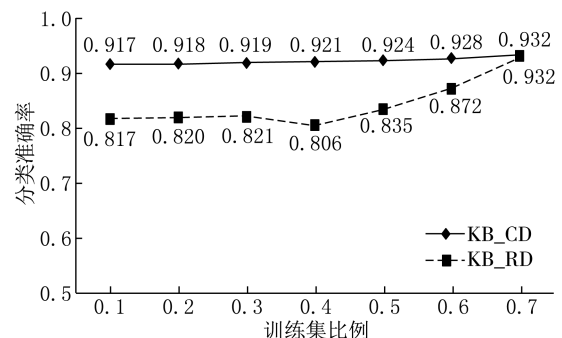


图 7 知识积累对谣言检测的影响

从图7得知,当争议检测任务只用了10%的训练集,谣言检测准确率达到0.917,比独立学习谣言检测任务时的分类准确率提升了0.2%。之后随着训练实例的增加,KB_CD逐渐累积,谣言检测效果呈现出越来越好的状态。当用到全部的争议检测训练集时,此时的准确率为0.932。同理,在谣言检测任务中,随着训练实例的增加,KB_RD逐渐累积,这时准确率先缓慢上升,当训练集比例为0.4时,谣言检测准确率有些降低,这是因为部分训练数据存在“噪声”,这将引起模型性能暂时下降。当模型积累的知识越来越多时,这些“噪声”所造成的影响力越来越小。当训练集比例从0.5开始增加时,谣言检测效果呈现出越来越好的状态,直至使用全部的谣言检测训练集时,准确率达到0.932。

3.3.4 模型分类效果对比

为了比较不同模型对谣言的检测效果,本文选取了SVM-TS^[12]、CAMI^[13]、EANN^[14]、GRU-2^[15]、textGCN^[16]、MHAN^[17]和BERT-BiGRU-MTL^[2]作为基线模型在相同的微博谣言数据集上展开实验。为保证实验的公平性,所有模型的训练集和测试集的比例保持一致。其中BERT-BiGRU-MTL属于多任务模型,在原文中使用情感分析作为辅助任务提升模型的谣言检测效果,且使用的情感分析数据较多,为了保证实验的公平性,将情感分析任务替换为本文中的争议检测任务,并得出实验结果。详细的实验对比结果如表5所示。

表5 模型分类效果对比

Model	A	F1
SVM-TS	0.631	0.630
CAMI	0.711	0.711
EANN	0.721	0.716
GRU-2	0.792	0.789
textGCN	0.857	0.856
MHAN	0.918	0.916
BERT-BiGRU-MTL	0.928	0.925
BERT-BiLSTM-LML	0.932	0.931

在表5中,SVM-TS属于传统机器学习方法,而CNN、EANN、GRU-2、textGCN和MHAN为深度学习方法。由于本文模型使用了BiLSTM提取的深层特征,并使用了争议检测作为辅助任务,所以模型的谣言检测效果很大程度上优于这些模型。而本文模型相比于多任务学习模型BERT-BiLSTM-LML,从评价指标上看虽然只有较少的提升,但多任务学习模型不能持续学习微博上不断变化的语言表达方式,也不能持

续提升模型的谣言检测效果。因此,无论是检测效果还是应用场景,本文提出的BERT-BiLSTM-LML都优于其他基线模型。

4 结束语

针对微博谣言带标签数据不足,且当下的谣言检测模型无法持续学习应对不断变化的微博网络语言等问题,本文提出BERT-BiLSTM-LML微博谣言持续检测模型,该模型以争议检测为辅助任务帮助模型提升谣言检测效果。实验结果表明:(1)模型在使用BERT词向量和BiLSTM提取的深层特征有效优化了模型效果;(2)辅助任务能有效提升谣言检测效果;(3)在持续学习过程中,随着知识的积累,模型在谣言检测任务上的学习效果也越来越好;(4)在各项评价指标上,本文所提出BERT-BiLSTM-LML模型都优于基线模型。尽管本文的工作提升了谣言检测效果,但在信息利用方面仍然具备一定的局限性。所以,接下来的工作将利用知识图谱进行增强实体语义信息,以达到更好的分类效果。

[参考文献]

- [1] 杨文太,梁刚,谢凯,等.基于突发话题和领域专家的微博谣言检测方法[J].计算机应用,2017,37(10):2799-2805.
- [2] 沈瑞琳,潘伟民,彭成,等.基于多任务学习的微博谣言检测方法[J].计算机工程与应用,2021,57(24):192-197.
- [3] MENDOZA M, POBLETE B, CASTILLO C. Twitter under crisis: Can we trust what we RT? [C] // Proceedings of the First Workshop on Social Media Analytics, July 25-28, 2010, Washington D. C., District of Columbia, New York: ACM, 2010: 71-79.
- [4] CHEN Z Y, HRUSCHKA E, LIU B. Lifelong machine learning and computer reading the web[J]. Synthesis Lecture on Artificial Intelligence and Machine Learning, 2016: 1-145.
- [5] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [6] CARUANA R. Learning many related tasks at the same time with backpropagation [C] // Proceedings of the 7th International Conference on Neural Information Processing Systems, 1 January 1994, Denver, Colorado, New York: ACM, 1994: 657-664.
- [7] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [EB/OL]. [2021-09-05]. 2018: arXiv: 1810.04805. <https://arxiv.org/abs/1810.04805>
- [8] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on

- sequence modeling [EB/OL]. [2021-09-05]. 2014; arXiv:1412.3555. <https://arxiv.org/abs/1412.3555>
- [9] PAUL RUVOLO, ERIC EATON. ELLA: An Efficient Lifelong Learning Algorithm[J]. ProcInt Conf Mach Learn, 2013:507-515.
- [10] ZHONG L, CAO J, SHENG Q, et al. Integrating semantic and structural information with graph convolutional network for controversy detection[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, Stroudsburg, PA, USA; Association for Computational Linguistics, 2020:515-526.
- [11] MA J, GAO W, WEI Z Y, et al. Detect rumors using time series of social context information on microblogging websites[C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, October 18-23, 2015, Melbourne, Australia. New York: ACM, 2015:1751-1754.
- [12] YANG F, LIU Y, YU X H, et al. Automatic detection of rumor on Sina Weibo [C] // Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, Beijing China, New York, NY, USA; ACM, 2012:1-7.
- [13] FENG YU, QIANG LIU, SHU WU, et al. A convolutional approach for misinformation identification[J]. Proceedings of the 26th International, 2017:3901-3907.
- [14] WANG Y Q, MA F L, JIN Z W, et al. EANN: Event adversarial neural networks for multi-modal fake news detection[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August 19-23, 2018, London, United Kingdom. New York: ACM, 2018:849-857.
- [15] MA JING, GAO WEI, MITRA PRASENJIT, et al. Detecting Rumors from Microblogs with Recurrent Neural Networks[C]//In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, New York: AAAI, 2016:3818-3824.
- [16] WANG Y Z, QIAN S S, HU J, et al. Fake news detection via knowledge-driven multimodal graph convolutional networks[C]//Proceedings of the 2020 International Conference on Multimedia Retrieval, June 8-11, 2020, Dublin, Ireland. New York: ACM, 2020:540-547.
- [17] 毕蓓, 潘慧瑶, 陈峰, 等. 基于异构图注意力网络的微博谣言监测模型[J/OL][2021-09-03]. <http://kns.cnki.net/kcms/detail/51.1307.tp.20210818.1855.004.html>.

Weibo rumors continuous detection model combining BERT word embedding and BiLSTM

HE Xiao-xia, GULANBAIER Tuerhong, MAIRIDAN Wushouer, WANG Song

(School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China)

Abstract: The field of microblog rumor detection currently uses unsupervised learning and multi-task learning methods to solve the problem of the lack of labeled data, but the former has poor stability, and the latter is difficult to continuously learn to cope with the constantly updated Internet language on Weibo. Lifelong Machine Learning paradigm aims to solve these problems, so a BERT-BiLSTM-LML model that continuously learns the two highly related tasks of controversy detection and rumor detection is proposed. First, the paper uses the BERT pre-training model to extract the word vectors of the two task input text data, and then uses the BiLSTM network to fully extract the text Contextual features. Finally, based on the BiLSTM deep features, the ELLA is used to model the feature data of the two tasks to achieve continuous detection of Weibo rumors. The experimental results show that the BERT word vector effectively optimizes the performance of the model. Compared with the Word2vec-BiLSTM-LML model based on the Word2vec word vector, the accuracy and F1 value are improved by 5.5%. Compared with independent learning, after continuing to learn the controversy detection task, the model's rumor detection accuracy increased by 1.7%, and the F1 value increased by 1.8%. At the same time, in the continuous learning process, with the accumulation of knowledge, the accuracy of rumor detection continues to increase. Finally, on the public Weibo data set, the BERT-BiLSTM-LML model rumor detection accuracy rate is 93.2%, and the F1 value is 93.1%, which is better than other baseline models.

Keywords: rumor detection; controversy detection; Lifelong Machine Learning; Weibo

(责任编辑:李亚军)