

文章编号:1007-5321(2020)05-0077-07

DOI:10.13190/j.jbupt.2020-074

一种用于图卷积网络的社交关系方向门控算法

李 蕾, 谢 旻, 蒋亚飞, 刘咏彬

(北京邮电大学 人工智能学院, 北京 100876)

摘要: 针对社交网络用户态度分析任务中用户之间原有社交关系方向可能阻碍态度信息流动以及标签扩散的问题,提出了一种应用于半监督图卷积网络的社交关系方向门控算法. 该算法首先在原有与逆向社交关系方向上分别进行图卷积运算,得到2种用户节点态度特征向量,然后利用门控机制对2种特征向量进行动态融合. 扩展了态度信息传播路径的同时,还能够捕捉用户影响力差异,以自动选择态度信息的流动方向. 在2个真实热点话题数据集上的实验结果表明,现有图卷积网络在加入该算法之后,其用户态度分析的准确率能够得到有效提升.

关键词: 图卷积网络; 门控算法; 用户态度分析; 社交网络

中图分类号: TP391

文献标志码: A

A Social Relationship Direction Gating Algorithm for Graph Convolutional Networks

LI Lei, XIE Yang, JIANG Ya-fei, LIU Yong-bin

(School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Facing the problem in social user attitude analysis that the original social relationship direction between users in social networks hinders the flow of attitude information and label propagation, a social relationship direction gating algorithm for graph convolutional networks is proposed. The algorithm first performs graph convolution on the origin and reverse social relationship directions to obtain two types of user node attitude feature vectors, and then leverages the gating mechanism to integrate the feature vectors dynamically. While expanding the propagation of attitude information, the algorithm can also capture the differences of user influence to automatically select the flow of attitude information. Experiments on two real hot topic datasets show that the accuracy of the existing graph convolutional networks can be effectively improved after adding the proposed algorithm.

Key words: graph convolutional network; gating algorithm; user attitude analysis; social network

随着计算机技术和互联网的不断发展,人们正处于一个以用户为中心,提倡“互动、分享、关系”的Web2.0时代. 社交网络为用户提供了一个表达和交换观点、获取和传播信息的平台,其热搜功能展示了时下人们正在讨论的话题. 当某一话题突然受到关注并成为社会热点时,及时分析用户的态度变得

尤为关键,对政府部门了解民情民生,第一时间掌握舆论动向,强化社会和谐关系具有重要意义.

社交网络具有天然图结构属性,用户可视为图中节点,用户之间相互关注这类社交关系可抽象为节点之间的边. 由于社交网络具备图的固有特性,加之大量数据导致人工标注成本高,基于图的半监

收稿日期: 2020-06-25

基金项目: 国家自然科学基金项目(91546121,71231002); 北京市科学技术委员会项目(Z181100001018035)

作者简介: 李 蕾(1974—),女,副教授,硕士生导师, E-mail:leili@bupt.edu.cn.

督学习方法被广泛应用于社交网络领域的研究。同时,随着深度学习的发展,图卷积网络(GCNs, graph convolutional networks)被广泛应用于图结构数据处理,极大地提升了基于图的半监督学习的性能。

社交网络如图1所示,节点 u_1 代表影响力较大且粉丝数量众多的用户,例如认证用户、名人和媒体账号等,节点 $u_2 \sim u_4$ 代表普通用户,实线箭头代表社交关系,即出点表示关注者,入点表示被关注者。虽然现有的GCNs能够很好地直接表征和处理有向图或无向图结构数据,但忽视了社交网络这类图结构数据中节点之间边的方向可能对态度信息传播以及态度标签扩散带来的影响,有如下3种情形:1) 受限于原有社交关系方向,影响力较大的节点 u_1 的态度信息无法向节点 u_3 和 u_4 逆向传播;2) 虽然节点 u_1 与节点 u_2 相互关注,但由于节点 u_1 具有更大的影响力,因此二者之间的态度信息流动方向更有可能是从节点 u_1 流向节点 u_2 ;3) 由于边的方向决定了标签扩散的路径,若 u_1 是已标记节点,则原有社交关系方向使得标签无法从 u_1 向 u_4 扩散。由此可见,虽然社交网络中的无向边能够使得态度标签在节点间自由扩散,但模型无法辨别关注者与被关注者以及两者之间的影响力差异,进而无法推测出2个用户节点之间态度信息如何传播;其次,虽然在直观上社交关系的逆向代表态度信息的流动方向,但这种态度信息的流动更多地发生在认证用户与粉丝之间,在数量更为庞大的普通用户之间,这种影响并不是绝对的。

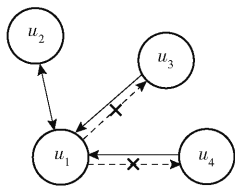


图1 社交网络示意图

针对以上问题,提出了一种应用于半监督图卷积网络的社交关系方向门控算法(SRDGA, social relationship direction gating algorithm),该算法能够灵活地与现有GCNs进行融合,提升其在社交网络中针对用户态度进行分类的准确率。自新浪微博采集了2个具有代表性的热点话题数据集,并通过大量的实验验证了SRDGA的有效性。与不加入SRDGA的GCNs相比,态度分类准确率平均提升了3.3%。数据集以及算法源码已公开^①。

1 相关工作

态度分析可以视为情感分析的子任务^[1],重点关注人们对某事或某人的喜欢或不喜欢的评价性反应^[2]。基于规则的情感分析方法需要人工整理情感词词典,例如程度副词表、否定词表以及表情符号的褒贬分类,可通过计算情感词语分数来判断文本的情感倾向^[3]。其缺点在于规则难以维护,需要不断更新规则来支持新的表达式和词汇表,且语言的灵活性和不合语法的现象会带来巨大的挑战。基于机器学习的情感分析方法利用机器学习方法^[4]对显式构建的情感特征进行学习。随着深度学习的发展,越来越多的基于神经网络的方法被提出^[5-6],这些方法能够自动学习文本的情感特征,同时可提升性能。

除此之外,由于社交网络存在天然的图结构,一些工作通过利用社交网络拓扑结构以辅助分析社交网络中的各类任务^[7-9]。但社交网络海量的数据导致有监督学习的成本很高,需要较多已标记训练数据才能实现较好的训练效果。基于图的半监督学习方法利用数据的图结构,可使模型只需很少的标签便可完成学习^[10]。此类方法大都使用聚类假设作为前提,即图中邻近节点倾向于共享相同的标签^[11]。但上述基于图的半监督学习方法仅利用了图结构本身,忽略了节点本身的特征。近年来,GCNs的出现和发展使得基于图的半监督学习性能有了极大的提升^[12-14]。Li等^[15]证明了GCNs是拉普拉斯平滑的一种特殊形式,平滑操作使得图中邻近的节点特征相似,从而简化了分类任务。然而,已有研究忽略了社交网络中用户节点的影响力差异对态度特征传播的影响,GCNs无法有效建模用户节点之间态度信息的流动。

2 SRDGA

2.1 模型框架

面向社交网络热点话题的用户态度分析模型由文本编码和图卷积2个阶段组成,如图2所示。在第1阶段中,用户所发表的微博文本由词向量矩阵表示,通过编码器对文本进行编码,抽取文本态度特征,生成整个微博的态度特征向量,并作为GCNs用户节点特征向量的初始化表示。在第2阶

① <https://github.com/xieyangbupt/SRDGA>

段中,GCNs 将处理整个社交网络图结构数据,通过卷积操作聚合节点的邻域特征,并更新当前节点特征,进行半监督学习,最终实现用户节点的态度分类任务,其中 SRDGA 将应用于图卷积阶段。

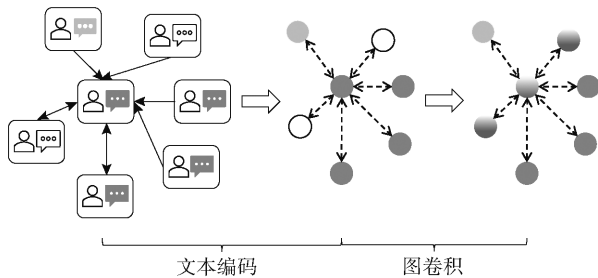


图2 态度分析模型框架

2.2 SRDGA 描述

由于社交网络中原有社交关系方向可能阻碍态度信息从影响力较大用户向普通用户的流动以及态度标签的扩散与传播,同时,无向社交关系无法判断关注者与被关注者以及二者的影响力差异。因此,基于门控机制的特点,可以学习并控制信息的流动以及多个信息源的取舍,故提出了可以灵活地与现有 GCNs 模型融合的 SRDGA,使得 GCNs 考虑用户影响力差异对态度传播带来的影响,能更有效地学习节点之间的信息流动,并扩展标签传播的路径。

定义社交网络图由 $G = (V, E, X)$ 表示,其中 V 为 $|V| = n$ 的用户节点集合; E 为社交关系集合,其中 $e_{ij} \in E$ 表示用户 i 关注了用户 j ; $X = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^{n \times c}$ 为节点特征矩阵, $x_i \in \mathbb{R}^c$ 表示用户 i 的微博特征向量,其中 c 表示特征向量的维度。已知态度标签的用户节点集合为 V_l ,目标是预测其余节点 V_u 的态度标签。

如图3和算法1所示,首先从原有社交关系 E 生成邻接矩阵 $A = [a_{ij}] \in \mathbb{R}^{n \times n}$,如果 $e_{ij} \in E$ 或 $i = j$,则 $a_{ij} = 1$;否则 $a_{ij} = 0$ 。然后对原有社交关系的方向进行反转,得到 $\bar{E} = \{e_{ji} | e_{ij} \in E\}$,同时生成邻接矩阵 $\bar{A} = [\bar{a}_{ij}] \in \mathbb{R}^{n \times n}$,如果 $a_{ij} = 1$,则 $\bar{a}_{ji} = 1$ 。之后,在第 k 层的 A 和 \bar{A} 上进行图卷积操作:

$$O^{(k)} = \sigma[f(X^{(k)}, \Theta^{(k)}, A)] \quad (1)$$

$$\bar{O}^{(k)} = \sigma[f(X^{(k)}, \Theta^{(k)}, \bar{A})] \quad (2)$$

其中: f 代表图卷积函数,它可以是图卷积网络(GCN, graph convolutional network)^[12]、采样与聚合(SAGE, sample and aggregate)^[13]和图注意力网络(GAT, graph attention network)^[14]等; σ 为激活函数; $X^{(k)}$ 为第 k 层的节点特征矩阵; $\Theta^{(k)}$ 为第 k 层

GCNs 的参数。之后,对 GCNs 在 A 和 \bar{A} 上得到的结果 $O^{(k)}$ 和 $\bar{O}^{(k)}$ 进行门控运算:

$$G^{(k)} = \text{sigmoid}(W_1^{(k)} O^{(k)} + W_2^{(k)} \bar{O}^{(k)} + b^{(k)}) \quad (3)$$

将 $O^{(k)}$ 和 $\bar{O}^{(k)}$ 通过门控结果 $G^{(k)}$ 进行融合,如式(4)所示,得到了下一层 GCNs 的节点特征矩阵输入 $X^{(k+1)}$ 。这样得到的 $X^{(k+1)}$ 融合了原有社交关系方向和逆向 2 个方向传播的节点特征,有

$$X^{(k+1)} = G^{(k)} \odot O^{(k)} + (1 - G^{(k)}) \odot \bar{O}^{(k)} \quad (4)$$

假设 GCNs 有 K 层,则把 GCNs 的第 K 层结果链接一个全连接层,使用 softmax 计算在 3 个态度类别上的概率,最终使用交叉熵作为模型的损失函数。

$$Z = \text{softmax}(WX^{(K)} + b) \quad (5)$$

$$L = - \sum_{i \in V_l} \sum_{j=1}^3 Y_{ij} \ln Z_{ij} \quad (6)$$

其中: V_l 为有标签顶点的集合, $Y \in \mathbb{R}^{|V_l| \times 3}$ 为有标签顶点的 one-hot 标签矩阵。

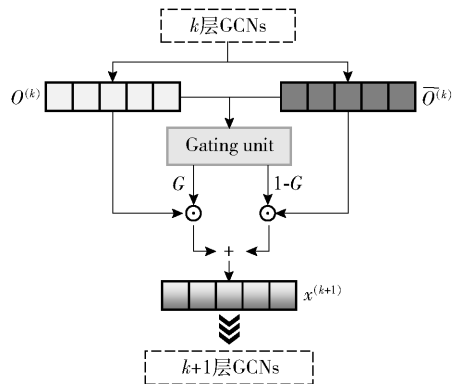


图3 SRDGA 图示

算法1 SRDGA

输入: GCN f and parameter $\Theta^{(i)}$, GCN layers k , initial node feature matrix $X^{(0)}$, adjacency matrix A and \bar{A} ;

输出: node feature matrix of last layer $X^{(k)}$

for $i \leftarrow 0$ to $k-1$ do

$$O^{(k)} = \sigma(f(X^{(k)}, \Theta^{(k)}, A))$$

$$\bar{O}^{(k)} = \sigma(f(X^{(k)}, \Theta^{(k)}, \bar{A}))$$

$$G^{(k)} = \text{sigmoid}(W_1^{(k)} O^{(k)} + W_2^{(k)} \bar{O}^{(k)} + b^{(k)})$$

$$X^{(k+1)} = G^{(k)} \odot O^{(k)} + (1 - G^{(k)}) \odot \bar{O}^{(k)}$$

end

3 数据集

数据采集自中国最大的社交网站新浪微博^①,

① <https://weibo.com/>

包括 2019-11-01—2019-12-01 期间大量讨论的“90 后该不该攒钱来改变现状”的话题,以及 2020-01-01—2020-03-01 期间大量讨论的“停课不停学”2 个热点话题下的相关数据,所构建的数据集分别用 T1 和 T2 表示. 数据集包含该热点话题下的微博以及用户之间的社交关系. 之后为这 2 个热点话题分别构建了图,图中节点表示用户,节点特征由微博文本表示,边由用户之间的社交关系生成,即若用户 u_i 关注了用户 u_j ,则生成一条从 u_i 到 u_j 的边. 之后使用 jieba^① 对文本进行分词和词性标注,并删除了停用词.

对数据的划分参考了 Yang 等^[16]对引文网络的半监督数据集划分方法,对采集到的 2 个热点话题数据集进行了划分和人工标注. 具体操作如下:对每个态度类别,随机抽取了 20 个实例作为有标签训练数据,500 个实例作为测试数据,其余用作无标签训练数据. 样本标签共分 +1、-1 和 0 三类,分别表示支持或赞成话题,反对或不赞成话题以及中立或没有明显态度. 表 1 所示为数据集的统计信息. 由于大量未标记数据的存在,因此只统计了测试集中的标签不匹配率,即 2 个相邻节点标签不一致的比例,以抽样说明数据集中相邻用户的态度不一致的比例. 同时,此项指标也在一定程度上体现了基于图的半监督学习所面临的挑战,因为此类方法都遵从于聚类假设,即相邻的节点倾向于拥有相同的标签.

表 1 数据集统计信息						
话题	节点数量	边数量	平均词数	平均字符数	已标记数据比例/%	标签不匹配率/%
T1	2 123	1 615	44	75	2.8	44.4
T2	17 694	30 805	51	94	0.3	47.2

4 实验结果与分析

4.1 实验设置

在文本预处理阶段,随机初始化维度为 256 维的词向量,并对文本进行截断或补零以达到平均长度. 在文本编码器和 GCNs 中间,GCNs 内部不同卷积层之间,以及和 GCNs 和全连接层之间添加了 0.5 的 dropout^[12]. 使用 Aadm^[17]作为优化器进行训练,学习速率设置为 0.005,并进行 100 次迭代训练. 采用准确率——计算分类器正确分类的样本数与总样本数之比作为模型分类结果的评价

指标.

4.2 对比模型

为了验证 SRDGA 的有效性及其迁移拓展能力,选取了多种文本编码器与 GCNs 的组合进行实验. 在文本编码阶段,分别使用了 TextCNN^[18]、Bi-LSTM^[19]、FastText^[20] 以及 Transformer^[21] 作为编码器,其中 TextCNN 使用了大小分别为 3、4、5 的卷积核,每种大小的卷积核设置为 100 个,输出向量维度为 300 维;Bi-LSTM 单向隐含层大小设为 150 维,输出向量维度为 300 维;FastText 的隐含层大小设为 300 维,输出向量维度为 300 维;Transformer 的隐含层大小设置为 300 维,输出向量维度为 300 维. 在图卷积阶段,分别使用了 GCN、SAGE 和 GAT 作为图卷积层,维度均为 300 维. 为防止过拟合,3 种 GCNs 只叠加了 2 个卷积层.

为了证明 GCNs + SRDGA 的性能优于 GCNs 直接在原有社交关系、逆向社交关系以及无向社交关系构成的图结构上进行卷积运算的性能,设置了如下 3 种社交网络图结构,并将 GCNs 在这 3 种图上的运行效果与 GCNs + SRDGA 作比较.

1) 原有社交关系图 (Origin): 社交网络图中节点代表用户,用户节点之间的边由数据集中的社交关系直接生成.

2) 逆向社交关系图 (Reverse): 在 Origin 的基础上将所有边的方向转变为逆向.

3) 无向社交关系图 (Undirected): 将 Origin 与 Reverse 进行合并,即将每条边扩展为无向.

除此之外,另外 2 种非神经网络方法也被引入进行对比.

1) 标签传播 (LP, label propagation): 该方法以微博的 TF-IDF 作为节点特征表示,利用节点特征相似度构建边权重,实现半监督学习.

2) 投票方法 (Vote): 针对某一用户节点,根据其邻域节点的态度标签,依据少数服从多数的投票原则决定其态度标签. 由于本实验数据含有大量未标记数据,所以整体采用迭代的方法,从已标记数据开始扩散态度标签,当已标记节点集合不再变化时,算法达到收敛.

4.3 实验结果及分析

2 个热点话题数据集 T1 和 T2 上的实验结果分别如表 2 和表 3 所示. 实验结果表明,GCNs +

① <https://github.com/fxsjy/jieba>

SRDGA 的准确率相较于 GCNs 在 Origin、Reverse 和 Undirected 均有提升. 其中 SRDGA 的准确率相较于 Origin 平均可提升 1.6%, 相较于 Reverse 平均可提升 5.3%, 相较于 Undirected 平均可提升 3.1%.

表 2 T1 数据集的实验结果

模型	Origin	Reverse	Undirected	SRDGA
TextCNN + GAT	60.4	54.0	57.6	63.0
TextCNN + GCN	53.2	54.8	56.2	56.4
TextCNN + SAGE	60.2	51.8	57.4	61.0
Bi-LSTM + GAT	55.8	43.4	52.4	56.4
Bi-LSTM + GCN	55.8	50.0	53.0	57.0
Bi-LSTM + SAGE	57.6	41.2	54.2	58.0
FastText + GAT	51.4	44.6	48.8	52.4
FastText + GCN	48.2	49.0	52.0	53.2
FastText + SAGE	52.6	46.6	52.8	54.2
Transformer + GAT	47.8	47.6	47.6	50.0
Transformer + GCN	47.0	47.4	47.6	48.0
Transformer + SAGE	47.2	47.6	48.0	49.0
LP	45.4	—	—	—
Vote	29.8	—	—	—

表 3 T2 数据集的实验结果

模型	Origin	Reverse	Undirected	SRDGA
TextCNN + GAT	50.0	47.6	47.4	51.4
TextCNN + GCN	49.6	50.0	46.6	50.2
TextCNN + SAGE	48.8	46.8	45.8	51.0
Bi-LSTM + GAT	43.8	40.0	39.6	44.6
Bi-LSTM + GCN	42.8	43.8	42.6	46.0
Bi-LSTM + SAGE	43.8	40.2	41.4	45.6
FastText + GAT	43.2	37.6	40.0	44.0
FastText + GCN	41.2	44.0	43.0	44.8
FastText + SAGE	44.9	41.6	45.0	45.0
LP	47.8	—	—	—
Vote	35.0	—	—	—

对比同一种文本编码器和 GCNs 的组合在不同社交关系图 Origin、Reverse 和 Undirected 上的实验结果, 可以发现, GCNs 在不同社交关系方向设置下的态度分类表现是存在差异的, 证实了边的方向对态度传播的重要性. 对比同一种组合在 T1 和 T2 上的实验结果以及同一数据集上不同组合之间的实验结果不难发现, 在 Origin、Reverse 和 Undirected 上的实验结果并不一定能保持一致性, 简单地将原有社

交关系转变为逆向, 或扩展为无向之后并不一定能提高模型的准确率, 而是与社交网络规模、所选取的文本编码器以及图卷积网络有关. 例如, 在 T1 和 T2 上, 模型在 Origin 上的性能较 Reverse 和 Undirected 普遍较高; T1 中, 模型在 Undirected 上的性能普遍较 Reverse 高, 而在 T2 中, Reverse 上的准确率普遍较 Undirected 上的高; 当使用 GCN 作为卷积层时, Reverse 的准确率较 Origin 高. 而再加入 SRDGA 之后, 模型在 2 个数据集上的准确率均得到了不同程度的提升, 进一步证明了 SRDGA 算法的有效性.

当 TextCNN 作为文本编码器时, 其态度分类准确率最高, 在 T1 上最高可达 63.0%, 在 T2 上最高可达 51.4%. 分类准确率接近甚至超过了标签匹配率, 可以说明新方法的有效性. Bi-LSTM 在 T1 上性能略逊于 TextCNN, 但在 T2 上与 TextCNN 存在差距, 其原因在于 T2 的平均文本长度相比于 T1 长, 同时态度的表达更能由局部语义凸显, 这也是 TextCNN 局部卷积操作能够得到更高准确率的原因, 而 Bi-LSTM 可能在关注长期语义依赖以及逆向语义特征时引入了额外的噪声, 导致性能下降. FastText 的优点在于参数较少同时结构简单, 但同时也导致捕捉态度特征的能力不足, 所以准确率较 TextCNN 和 Bi-LSTM 有所降低. 由于 Transformer 其复杂且强大的注意力机制以及解决长距离依赖问题的能力在机器翻译、自动文摘以及文本分类领域取得了优异的成绩. 然而在 T1 的实验中其准确率相较于 TextCNN、Bi-LSTM 以及 FastText 却低了最高 13% 左右. 其原因在于 Transformer 作为一种较为复杂的监督学习模型, 需要大量的标记数据才能进行充分的训练, 而社交网络中采用的是半监督学习, 其有标记数据较少, 因此对复杂模型学习是一项挑战, 从而导致其准确率不高. 且由于硬件条件所限, 未在 T2 上对其进行测试. 但根据其在较小规模数据集 T1 上的表现, 可以合理推测在有标记数据占比更少的 T2 上的效果也将不如其他 3 种文本编码器.

在同一种文本编码器的条件下, GAT 与 SAGE 作为图卷积层的性能普遍较 GCN 的性能高, 其原因在于 GAT 和 SAGE 均首先将邻域向量聚合成为一个向量, 然后再通过聚合节点向量和邻域向量来完成节点隐含表示的更新. 因此, GAT 和 SAGE 可以清楚地区分用户自身与关注用户或粉丝, 在聚合时使得模型自动学习如何融合自身态度特征与邻域态度特征, 比 GCN 更具灵活性和通用性.

LP 在 T1 上的表现比所有深度神经网络模型均有所欠缺,准确率较 SRDGA 的平均准确率低 9.4%,原因在于 TF-IDF 特征比较浅层,无法准确表达态度信息,而且 LP 方法使用文本相似性构建节点之间的边,在一定程度上丢失了原有社交结构信息。但 LP 在较大规模数据集 T2 上的准确率超过了 T1 上的准确率,且较 SRDGA 的平均准确率高 1.1%,超过了部分深度神经网络模型,如 Bi-LSTM 与 FastText 作为文本编码器的某些组合。其原因在于 TF-IDF 特征属于统计意义上的特征,故如果数据规模越大,那么 TF-IDF 值越准确,并更接近语料中词语的真实分布,从而使得准确率有所提高;但是相比于某些深度神经网络能够挖掘的特征,其特征还是比较浅层,因此 LP 性能无法达到最优。另一方面,基于深度神经网络的半监督方法需要部分有标记数据提供监督信号,所以随着数据规模的扩大,在已标记数据占比越来越低的情况下,其准确率也会有所下降。

Vote 方法在 2 个数据集上的表现都较差,准确率较 SRDGA 的平均准确率分别低 25% 和 11.6%,原因在于数据集中的已标记数据占比较低,绝大部分都是未标记数据。因此,未标记节点周围的已标记数据较为稀疏,投票方法无法发挥最大作用。除此之外,数据中还存在部分孤立节点,Vote 方法无法有效处理此类情况,而基于神经网络的方法本质上学习到了节点特征到态度标签的映射关系,因此也可以较好地对孤立节点进行分类。

综上所述,GCNs 在与 SRDGA 算法融合之后的模型性能均优于 GCNs 在 Origin、Reverse 和 Undirected 图上的结果,且优于非神经网络方法 LP 和 Vote,这都很好地证明了 SRDGA 算法的提出和使用能够较好的弥补 GCNs 应用于社交网络图结构的不足,在多种情况下都能有助于提升系统的整体性能。

5 结束语

社交关系方向门控算法 SRDGA 可灵活加入现有 GCNs,利用门控机制自动学习如何融合在原有社交关系以及逆向社交关系卷积得到的节点特征,使得 GCNs 在捕捉不同影响力用户之间态度特征的传递的同时,扩展了标签传播的路径。在 2 个真实数据集上进行的多种文本编码器与 GCNs 的组合实验证明了 SRDGA 的有效性,提升了现有 GCNs 态度分类的准确率。未来的工作将在更多数据集以及不同

的已标记数据比例上进行更为广泛的实验,一方面尝试利用用户的属性信息扩展社交网络拓扑结构;另一方面尝试将深度神经网络模型与 LP 模型进行融合,不断提升态度分析的性能;也将尝试将算法应用于社交网络中的情感分析,以验证算法在不同任务上的泛化能力。

致谢 本研究工作得到了教育部信息网络工程研究中心的支持。

参考文献:

- [1] Bing Liu. Sentiment analysis and opinion mining [J]. Synthesis Lectures on Human Language Technologies, 2012, 5(1): 1-167. [LinkOut]
- [2] Myers D G. Social psychology [M]. New York: Mcgraw-Hill, 2009: 124.
- [3] 李勇敢,周学广,孙艳,等. 中文微博情感分析研究与实现 [J]. 软件学报, 2017, 28(12): 3183-3205.
Li Yonggan, Zhou Xueguang, Sun Yan, et al. Research and implementation of Chinese microblog sentiment classification [J]. Journal of Software, 2017, 28(12): 3183-3205.
- [4] 李婷婷,姬东鸿. 基于 SVM 和 CRF 多特征组合的微博情感分析 [J]. 计算机应用研究, 2015, 32(4): 978-981.
Li Tingting, Ji Donghong. Sentiment analysis of microblog based on SVM and CRF using various combinations of features [J]. Application Research of Computers, 2015, 32(4): 978-981.
- [5] 周瑛,刘越,蔡俊. 基于注意力机制的微博情感分析 [J]. 情报理论与实践, 2018, 41(3): 89-94.
Zhou Ying, Liu Yue, Cai Jun. Sentiment analysis of micro-blogs based on attention mechanism [J]. Information Studies (Theory & Application), 2018, 41(3): 89-94.
- [6] 杜慧,俞晓明,刘悦,等. 融合词性和注意力的卷积神经网络对象级情感分类方法 [J]. 模式识别与人工智能, 2018, 31(12): 1120-1126.
Du Hui, Yu Xiaoming, Liu Yue, et al. CNN with part-of-speech and attention mechanism for targeted sentiment classification [J]. Pattern Recognition and Artificial Intelligence, 2018, 31(12): 1120-1126.
- [7] 孙晓,彭晓琪,胡敏,等. 基于多维扩展特征与深度学习的微博短文本情感分析 [J]. 电子与信息学报, 2017, 39(9): 2048-2055.
Sun Xiao, Peng Xiaoqi, Hu Min, et al. Extended multi-modality features and deep learning based microblog short text sentiment analysis [J]. Journal of Electronics & In-

- formation Technology, 2017, 39(9): 2048-2055.
- [8] West R, Paskov H S, Leskovec J, et al. Exploiting social network structure for person-to-person sentiment analysis[J]. Transactions of the Association for Computational Linguistics, 2014, 2: 297-310.
- [9] 黄丹阳, 王菲菲, 杨扬, 等. 基于网络结构与用户内容的动态兴趣识别方法[J]. 北京邮电大学学报, 2018, 41(2): 103-108.
- Huang Danyang, Wang Feifei, Yang Yang, et al. Dynamic interest identification based on social network structure and user generated contents[J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(2): 103-108.
- [10] Zhu X, Ghahramani Z, Lafferty J D. Semi-supervised learning using gaussian fields and harmonic functions [C]//Proceedings of the 20th International Conference on Machine Learning (ICML-03). Washington, DC: AAAI Press, 2003: 912-919.
- [11] Chapelle O, Zien A. Semi-supervised classification by low density separation[C]//AISTATS. 2005: 57-64.
- [12] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [EB/OL]. 2016 (2016-09-06) [2020-06-25]. <https://arxiv.org/abs/1609.02907>.
- [13] Hamilton W L, Ying R, Leskovec J. Inductive representation learning on large graphs [EB/OL]. 2017 (2017-06-07) [2020-06-25]. <https://arxiv.org/abs/1706.02216>.
- [14] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks [EB/OL]. 2017 (2017-10-30) [2020-06-25]. <https://arxiv.org/abs/1710.10903>.
- [15] Li Q, Han Z, Wu X M. Deeper insights into graph convolutional networks for semi-supervised learning [C]//Thirty-Second AAAI Conference on Artificial Intelligence. Palo Alto: Association for the Advancement of Artificial Intelligence, 2018: 3538-3545.
- [16] Yang Zhilin, Cohen W W, Salakhutdinov R. Revisiting semi-supervised learning with graph embeddings [EB/OL]. 2016 (2016-03-29) [2020-06-25]. <https://arxiv.org/abs/1603.08861>.
- [17] Kingma D P, Ba J. Adam: a method for stochastic optimization [EB/OL]. 2014 (2014-12-22) [2020-06-25]. <https://arxiv.org/abs/1412.6980>.
- [18] Kim Y. Convolutional neural networks for sentence classification [EB/OL]. 2014 (2014-08-25) [2020-06-25]. <https://arxiv.org/abs/1408.5882>.
- [19] Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging [EB/OL]. 2015 (2015-08-09) [2020-06-25]. <https://arxiv.org/abs/1508.01991>.
- [20] Joulin A, Grave E, Bojanowski P, et al. Bag of tricks for efficient text classification [EB/OL]. 2016 (2016-07-06) [2020-06-25]. <https://arxiv.org/abs/1607.01759>.
- [21] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems. Red Hook: Curran Associates Inc, 2017: 6000-6010.