

基于贝叶斯衍生分类器的社交网络用户影响力评价模型

周春良¹, 刘仰光¹, 孟祥佩²

(1. 宁波财经学院金融与信息学院, 浙江 宁波 315175; 2. 宁波财经学院基础学院, 浙江 宁波 315175)

摘要: 为了防止社交网络中的负面信息快速传播, 则需要通过评价社交网络中用户的影响力来找出影响力大的社交网络节点。针对传统算法在社交网络领域中交叉特性缺失的问题, 结合高斯贝叶斯衍生分类器, 提出一种网络用户影响力评价模型。该模型结合用户活跃度、用户联系度、用户覆盖度等维度, 建立社交网络用户影响力刻画指标, 同时考虑社交网络用户之间的关系特征和用户自身的行为特征, 降低僵尸粉和垃圾社交网络对网络评价结果的影响, 通过建立连续属性朴素贝叶斯分类器方法, 提出基于高斯贝叶斯衍生分类器的模型求解方法。使用新浪微博中 152 059 423 条媒体报纸用户评论作为实验数据, 分析影响该评价模型的关键因素, 利用仿真软件完成和 HRank 等传统模型对比实验, 结果表明, 该模型体现了社交网络用户的交叉特性, 提升了模型的实用性, 相比于其他传统算法, 该模型分类误差更趋于稳定, 分类结果的误差率更低, 适应性更好。

关键词: 社交网络; 影响力; 贝叶斯衍生分类器; 评价模型; 用户活跃度

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0066957

Social Network User Influence Evaluation Model Based on Bayesian Derived Classifier

ZHOU Chunliang¹, LIU Yangguan¹, MENG Xiangpei²

(1. College of Finance and Information, Ningbo University of Finance and Economics, Ningbo 315175, Zhejiang, China;

2. Academic College, Ningbo University of Finance and Economics, Ningbo 315175, Zhejiang, China)

【Abstract】 To prevent the rapid spread of negative information on social networks, influential social network nodes must be identified by evaluating the influence of users on social networks. To address the problem of missing cross-characteristics in traditional algorithms in the field of social networks, this study proposes a network user influence evaluation model combined with Gaussian Bayesian-derived classifiers. The model first combines user activity, contact, coverage, and other dimensions to establish social network user influence characterization indicators. Simultaneously, the relationship characteristics between social network users and user behavior characteristics are considered to reduce the impact of zombie fans and garbage social networks on network evaluation results. A model-solving method based on a Gaussian Bayesian classifier is proposed by establishing a continuous attribute-naive Bayes classifier method. The key factors affecting the evaluation model are analyzed in depth using 152 059 423 media and newspaper user comments from Sina Weibo as experimental data. A comparative experiment with traditional models, such as HRank, is conducted using simulation software to verify the feasibility of the model. The experimental results indicate that the model reflects the cross-characteristics of social network users and improves its practicality. Compared to traditional algorithms, this model tends to have more stable classification errors, lower error rates in the classification results, and better adaptability.

【Key words】 social network; influence; Bayesian derived classifier; evaluation model; user activity

0 引言

随着移动智能的不断普及, 微博、微信等社交媒体软件快速发展。大量社交用户之间的关注关系构成了一个常见的社会网络^[1-3]。其中社交网络用户影响力越大, 其对信息的传播作用也就越大^[4-6]。社交网络为人们日常生活提供便利^[7-9], 但

也给社会带来了负面影响, 比如在新型冠状病毒感染疫情防控事件中, 谣言利用社交网络进行传播, 信息传播速度快、范围广, 不利于对社交网络信息管理, 极易造成网络舆论, 其影响范围巨大, 传播速度较快。

为了减少社交网络对社会带来的负面影响, 防止谣言等信息在社交网络中快速传播, 需要在社交

收稿日期: 2023-02-16 修回日期: 2023-05-21

基金项目: 浙江省哲学社会科学规划课题(21NDJC167YB); 国家自然科学基金(62001199); 宁波财经学院科研基金硕士学位培育点项目(1320230911)。

通信作者 E-mail: zhouchunliang@nbufe.edu.cn

网络中寻找出影响力大的各类网络节点,对其进行监控。因此,研究社交网络用户影响力评价模型是十分必要的。

目前,社交网络影响力问题主要采用各类智能算法和模型解决^[10],包括 PageRank 算法、隐半马尔可夫模型、各类影响扩散模型以及行为动力学模型,但这些方法没有考虑社交网络领域的交叉特性,并且很多研究没有基于社交用户关系网络,这就大大降低了其结果的实用性和参考价值。

针对社交网络领域的交叉特性缺失的问题,本文结合高斯贝叶斯衍生分类器^[11-13],提出一种社交网络用户影响力评价模型。首先给出社交网络用户影响力刻画指标,同时针对社交网络用户之间的关系特征和用户自身的行为特征,利用高斯贝叶斯衍生分类器给出上述模型的求解方法。最后利用仿真实验研究影响该评价模型的关键因素。

1 相关研究

目前,评价和评估社交网络的模型主要有:1)基于行为动力学模型;2)行为扩散力模型;3)基于 TwitterRank、ObjectRank、隐半马尔可夫模型等其他方法。

1.1 行为动力学模型

文献[1]将用户行为定义为 3 种:跟踪,转发和提及,深入分析了这 3 种用户行为,并通过实验得出用户最关心的是具备大量信息来源的公共信息。文献[14]从人类动力学的角度来解释社交网络中多对多模式社交网络发布-评论行为的过程,解释了社交网络用户信息之间的交互关系,为社交网络舆情的监控与引导提供了一定的理论依据。文献[15]针对社交网络用户关系网络的内部特征,基于平均场理论提出一种社交网络用户关系网络演化模型,深入分析谣言在该演化模型上的传播动力学行为以及该演化模型的拓扑统计特性,结果表明社交网络用户关系网络具有无标度特性,并且度分布指数不仅与反向连接概率有关,而且还与节点吸引度分布相关。文献[16]提出一种基于网络用户交互的好友推荐算法,该算法与只考虑网络拓扑结构或用户属性的算法相比,通过实验验证了该算法的准确率更高。文献[17]通过分析用户之间的信任度和影响因素来量化用户的信任值,通过网络关系来建立信任传递模型,该模型为好友互动和好友建立提供了有效推荐。

1.2 影响扩散模型

影响扩散模型^[18]关注帖子的内容,根据帖子结

构和回复中的关键词,即关键词之间的转换关系计算参与者的活动。文献[19]提出了一种跨多社交网络的谣言传播模型,该模型能够模拟网络谣言的传播趋势,通过实验显示,该模型能够快速抑制谣言的传播。文献[20]基于社会影响力在社交网络的传播状况,结合社会网络结构和用户行为因素,提出了通过预测网络用户传播信息能力大小来度量用户社会影响力的方法,通过对比实验可以看出,该模型获得了更好的用户影响力估计结果。文献[21]提出一种基于遍历约束和交互信息的社交网络用户表征算法,对社交网络遍历进行约束以提高算法的时间效率,利用用户交互信息修改 Word2vec 模型,提高社交网络表征的准确率。

1.3 其他模型

文献[22]通过改进的 PageRank 算法,提出了一种 TwitterRank 算法,该算法用于测量主题敏感度推特的影响,从两个方面衡量了对主题的不利影响,考虑到用户之间的相似性和链接结构。文献[23]通过改进的 PageRank 算法,提出了一种 ObjectRank 算法,该算法用于根据链接评估用户在 Twitter 中的权限得分分析,为每个用户行为提供计算的初始权重。文献[24]提出了一种基于隐半马尔可夫模型的社交网络流行信息检测方法,利用该模型来刻画流行信息的传播过程,计算了每条信息在传播过程中产生的观测序列相对于模型的平均对数似然概率,并且实时更新了每条信息的流行度。文献[25]为了有效刻画社交网络的用户关系,将社交网络视作加权无向图,给出了基于各种用户属性信息的用户相似度计算方法。实验结果表明,基于社交信息的用户相似度在用户关系分析、团体挖掘方面具有较好适应性。

2 社交网络用户影响力评价模型

本节对社交网络用户的信息进行建模,构建了基于高斯贝叶斯衍生分类器的网络用户影响力模型,阐述了社交网络用户影响力的求解方法^[26]。

2.1 社交网络用户影响力刻画方法

微博、微信等作为一种便捷的社交平台对信息的传播起着重要作用,由于社交网络活跃用户数广,信息发布方式广,信息传播速度快、范围广,不利于对社交网络信息管理,极易造成网络舆论,由此本文建立一种社交网络用户的影响力评估体系,旨在通过实际的数据计算出用户的影响力指数,便于对社交网络社交平台进行管理。

用户影响力是衡量社交网络用户传播能力的重要指标,影响力越大,该用户的传播能力就越强,对

于整个社会的影响作用也大,用户的影响力因素包含用户的关注和粉丝关系以及用户的各类行为,这些因素将直接决定用户的影响力,通过对用户的影响力因素进行分析来构成影响力评估体系。本文采用信息传播实际影响人次覆盖率 H 、用户活跃度和用户联系度作为社交网络影响力的评估指标。覆盖率包括粉丝数量和关注数量,活跃度包括用户原创社交网络的数量、转发社交网络的次数、评论的次数以及私信的数量。传播力包含了所有原创社交网络被转发、被阅读、被收藏、被点赞、被评论的次数总和。

社交网络用户影响力评价具体算法如下:

算法 1 社交网络用户影响力评价算法

输入 网络用户数据流特征集

输出 网络用户的影响力排序

步骤 1 初始化数据,并对数据进行预处理。

步骤 2 根据用户原创社交网络数量、质量、知名度,计算网络用户自身的权威度。

步骤 3 剔除沉默粉用户,计算用户的活跃度。

步骤 4 根据网络与网络之间的连接关系,计算用户的联系度。

步骤 5 对每个用户的粉丝数量进行测度,计算用户的覆盖度。

步骤 6 根据用户活跃度、联系度和用户覆盖度,计算社交网络用户的影响力。

步骤 7 循环步骤 2~步骤 6,直到所有用户的影响力计算完成,跳转到步骤 8。

步骤 8 对用户进行影响力排序,能更加直观地找出影响力高的用户。

具体方法如下:

1)初始化数据。从所有社交网络用户中随机选取 N 个活跃度较高、联系度较大的用户,对这 N 个用户的社交网络数据进行采集,采集的数据内容包括原创社交网络数、粉丝数、关注的数目、被转发的数目、转发的数目、被评论的次数、评论的次数、被收藏的数目、收藏的数目等。

2)计算用户自身的权威度,包含用户原创社交网络的数量和质量、用户的知名度。其中用户的知名度会在一定程度上提升用户的影响力。知名度高的用户,其言论更容易产生影响。同时,社交网络的内容质量也是一个重要的影响因素,一篇高质量、观点独到的见解或者具备娱乐性的内容,容易吸引更多的关注并产生影响,会影响用户的权威度,其计算公式如下:

$$W(i) = \frac{\sum_{k=0}^n C(i) \times Q(i) \times Z(i)}{t} \quad (1)$$

式中: $W(i)$ 表示用户 i 的权威度; $C(i)$ 表示用户 i 的发帖数量; $Q(i)$ 表示用户 i 的发帖质量; $Z(i)$

表示用户 i 的知名度; n 为某一时间用户发帖次数; t 为某一时间段的时间间隔。

3)计算用户的活跃度。用户的活跃度能在一定程度上影响用户的影响力,在社交网络中存在很多的沉默粉,这些沉默粉虽然关注了某些用户,但其对消息的传播没有帮助,因而其对用户的影响力也无帮助,因此通过多用户的活跃度来剔除沉默粉,活跃度的计算公式如下:

$$T(i) = \frac{\omega_1 N + \omega_2 F + \omega_3 L}{t} \quad (2)$$

式中: $T(i)$ 表示用户 i 的活跃度; N 表示用户 i 原创社交网络的数目; F 表示用户 i 的粉丝数量; L 表示用户 i 的关注数; ω_1 、 ω_2 、 ω_3 分别为相应影响因素的权重值; t 为某一时间段的时间间隔。

4)计算用户的联系度。每个社交网络用户之间都有相互的联系,联系度越高,信息传播的能力就越大,进而也会提升用户的影响力,用户之间的联系度与用户的活跃度计算类似,其计算公式如下:

$$T(i, j) = \frac{\lambda_1 U + \lambda_2 V + \lambda_3 S}{t} \quad (3)$$

式中: $T(i, j)$ 表示用户 i 和用户 j 的联系度; U 表示用户 i 转发用户 j 的社交网络数目; V 表示用户 i 评论用户 j 的社交网络数目; S 表示用户 i 收藏用户 j 的社交网络数目; λ_1 、 λ_2 、 λ_3 分别为相应影响因素的权重值; t 为某一时间段的时间间隔。

5)计算用户覆盖度。用户覆盖度即为用户 i 的社交网络的活跃粉丝数,其对社交网络的传播能力有很大影响,进而能间接提高社交网络用户的影响力,其计算公式如下:

$$H(i) = \frac{\sum R \cup V \cup M}{N} \quad (4)$$

式中: R 、 V 和 M 分别为对用户转发、评论、提及进行统计时所覆盖的人群; N 为网络的所有节点个数。

6)计算社交网络用户的影响力。本文所提的影响力模型由用户活跃度、用户联系度、用户覆盖度三大指标构成,其计算公式如下:

$$X = (\alpha T(i) + \beta T(i, j) + \gamma H(i)) \times W(i) \quad (5)$$

式中: $T(i)$ 表示用户 i 的活跃度; $T(i, j)$ 表示用户 i 和用户 j 的联系度; $H(i)$ 表示用户 i 的覆盖度; $W(i)$ 为用户 i 的权威度; α 、 β 和 γ 分别为 3 种影响因素的权重值。

7)循环第 2)步~第 6)步,直到所有用户的影响力计算完成,跳转到第 8 步。

8)对用户的影响力进行排序。对用户进行影响力排序,能更加直观地找出影响力高的用户。

2.2 基于高斯贝叶斯衍生分类器的用户影响力求解方法

高质量的社交网络粉丝通常会更积极地转发和分享社交网络内容,从而帮助吸引更多的关注和粉丝。扩大用户的影响力,参与社交网络的内容和讨论,主动留言、点赞、评论并与用户进行交流,这种互动能够提高社交网络的可见度和吸引力。但是,社交网络中也存在大量的僵尸粉,进行虚假的点赞和评论,误导其他用户,让用户对某些内容产生过高的期望或者认同。

为了减少僵尸粉和垃圾社交网络对评价结果的影响,本文提出了基于高斯贝叶斯衍生分类器的社交网络用户影响力评价模型,该模型既考虑各个社交网络用户之间的关系特征,又考虑了各用户自身的行为特征。通过建立连续属性朴素贝叶斯分类器,并将其用于对僵尸粉的鉴别,能在一定程度上提高分类效率和可靠性。其具体算法如下:

算法 2 基于高斯贝叶斯衍生分类器的用户影响力求解算法

输入 网络用户数据流特征集

输出 网络用户的影响力排序

步骤 1 初始化数据,并对数据进行预处理。

步骤 2 根据社交网络属性求解社交网络数据集的记录观测值。

步骤 3 通过高斯贝叶斯衍生分类器对用户进行分类。

步骤 4 对选取的每个用户的粉丝数和真实粉丝数进行相似度测量,剔除僵尸粉用户。

步骤 5 根据式(2)~式(4)分别计算用户的活跃度、联系度及覆盖度。

步骤 6 根据用户活跃度、联系度和用户覆盖度计算社交网络用户的影响力。

步骤 7 对用户进行影响力排序,能更加直观地找出影响力高的用户。

具体方法如下:

1)初始化数据。从所有社交网络用户中随机选取活跃度较高、联系度较大的用户,并对这些用户的社交网络数据进行采集。采集的数据内容包括原创社交网络数、粉丝数、关注的数目、被转发的数目、转发的数目、被评论的次数、评论的次数、被收藏的数目、收藏的数目等。

2)用 $X_1, X_2, \dots, X_n, C$ 表示连续社交网络属性(原创社交网络数、粉丝数、关注的数目、被转发的数目、转发的数目、被评论的次数、评论的次数、被收藏的数目、收藏的数目等)和类,即 $x_1, x_2, \dots, x_n, c$ 为其值, D 为选取的具有 N 个记录的数据集,数据随机产生于混合分布 P ,其中 $x_{ij} (1 \leq i \leq n, 1 \leq j \leq$

$N)$ 和 c_j 表示 X_i 和 C 在数据集 D 中的第 j 个记录观测值。

3)通过高斯贝叶斯衍生分类器将用户进行分类,剔除僵尸粉用户。假设 G_1, G_2 是两个 k 维总体, G_1 为真实粉用户, G_2 为僵尸粉用户,其分布密度分别为 $p_1(u), p_2(u)$, 用户群 $u = (u_1, u_2, \dots, u_m)$, 其中 u_i 来自 G_1 的概率为 q_1 , 来自 G_2 的概率为 $(1-q_1)$, 将 k 维空间 R^k 分为 (R^1, R^2) , 满足:

$$R^k = \{R_1, R_2 \mid R_1 \cup R_2 = R^k, R_1 \cap R_2 = \emptyset\} \quad (5)$$

若 $u \in R^1$, 则表示用户 u_i 来自真实用户群体 G_1 , 若 $u \in R^2$, 则表示用户 u_i 来自僵尸用户群体 G_2 。任何分类方法都不能满足 100% 的正确率, 而按此分类所得到的将真实用户群体 G_1 误判为僵尸用户群体 G_2 的概率计算公式如式(6)所示, 将僵尸用户群体 G_2 误判为真实用户群体 G_1 的概率计算公式如式(7)所示:

$$p(2 \mid 1, R) = \int_{R_2} p_1(u_i) dx \quad (6)$$

$$p(1 \mid 2, R) = \int_{R_1} p_2(u_i) dx \quad (7)$$

分类结果的平均分类误差 $f(R_1, R_2)$ 计算公式如下:

$$f(R_1, R_2) = q_1 c(2 \mid 1) P(2 \mid 1, R) + q_2 c(1 \mid 2) P(1 \mid 2, R) \quad (8)$$

4)对选取的每个用户的粉丝数 F_i 和真实粉丝数 F'_i 进行相似度测量, 计算公式如下:

$$R = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{F_i - \bar{F}}{S_{F_i}} \right) \cdot \left(\frac{F'_i - \bar{F}'}{S_{F'_i}} \right) \quad (9)$$

$$\bar{F} = \frac{1}{n} \sum_{i=1}^n F_i \quad (10)$$

$$\bar{F}' = \frac{1}{n} \sum_{i=1}^n F'_i \quad (11)$$

$$S_{F_i} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (F_i - \bar{F})^2} \quad (12)$$

$$S_{F'_i} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (F'_i - \bar{F}')^2} \quad (13)$$

式中: F 为用户粉丝数; $\bar{F}_i$ 为用户真实粉丝数均值; $S_{F_i}, S_{F'_i}$ 分别为用户粉丝均值和用户真实粉丝均值, 通过两者对比, 便能对粉丝列表进行真实还原, 剔除用户粉丝列表中的僵尸粉, 然后对剔除僵尸粉后的数据进行重新统计。

使用带标签的样本数据来训练贝叶斯分类器, 标签可以根据已知的真实用户和僵尸粉进行标记, 根据已有的数据和知识用于描述真实用户和僵尸粉

的分布情况。

5)根据式(2)~式(4)分别计算用户的活跃度、联系度及覆盖度。

6)根据式(5)计算各用户的影响力。

7)对各用户影响力按由大到小的顺序进行排序。

3 仿真实验

本文以 2021 年 10 月采集的新浪微博用户作为实验数据,统计包括用户的基本信息、粉丝列表及用户在近 30 天内所发微博的 152 059 423 条数据,将本文所提的模型以及 HRank 模型算出的用户影响力排名与新浪微博现有的用户影响力排名进行对比,并对各模型的用户影响力前 10 名进行统计,得到结果如表 1 所示。

表 1 各模型下媒体报纸影响力排名

Table 1 Ranking of media and newspaper influence under each models

排名	新浪微博	微博数/条	HRank 模型 排名	本文模型 排名
1	《人民日报》	50 738 363	2	1
2	《环球时报》	6 842 857	1	2
3	《楚天都市报》	7 549 425	3	3
4	《中国青年报》	4 530 902	4	4
5	《新民晚报》	5 073 876	7	6
6	《新京报》	23 394 642	6	5
7	《羊城晚报》	10 646 399	8	8
8	《新闻晨报》	29 115 536	5	7
9	《半岛都市报》	6 177 043	9	9
10	《大河报》	7 990 380	10	10

表 1 是对微博媒体势力榜的报纸类微博用户进行统计计算后得到的各模型下用户影响力前 10 名的用户,从表 1 反映出粉丝数量多的用户不一定影响力大,并且从表中可以看出本文模型计算出的用户影响力排名和新浪微博的用户影响力排名基本相同,由此可以看出本文所提模型是可行的。

图 1 描述的是对不同算法下排名前 $p\%$ 的用户

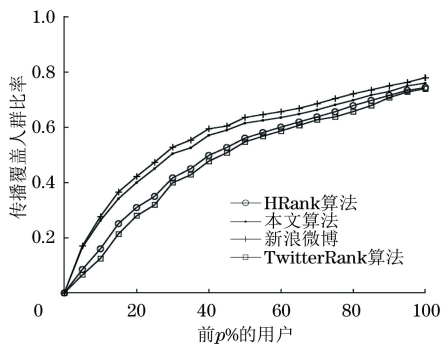


图 1 不同算法覆盖率对比

Fig.1 Comparison of coverage rate of different algorithms

影响人次覆盖率进行比对,考察各算法的优劣关系,从图 1 可以看出,本文所提的算法效果略优于 HRank 算法和 TwitterRank 算法,虽比新浪微博的真实排名略低,但却更接近真实排名,这说明本文算法对微博信息的分类和筛选存在合理性。

图 2 是用户日均微博数的分布,通过对用户日均微博数进行统计,能更直观地剔除这部分用户中不活跃用户或僵尸用户。从图 2 可以看出,大部分的用户日均微博数在 2~3 条,则可将日均微博数少于 1 条或只发 1 条的用户视为不活跃用户或僵尸用户,在进行用户影响力评价时可以不考虑该类用户。

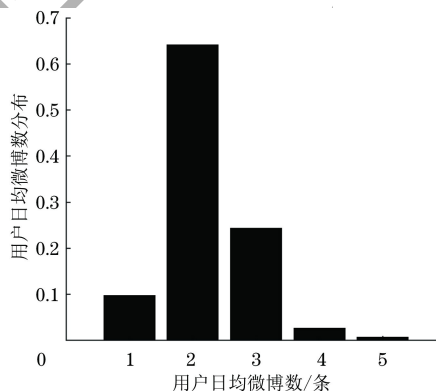


图 2 用户日均微博量分布

Fig.2 Average daily Weibo volume distribution of users

前文提到用户的活跃度对用户的影响力排名有关系,本文对用户活跃度与用户影响力的关系进行研究,结果如图 3 所示。从图 3 可以看出,用户活跃度在一定范围内与用户影响力排名几乎呈线性关系,但超过某一范围后用户的活跃度就对用户的影响力排名几乎没太大的影响,这就说明用户的影响力排名还受其他因素的影响。

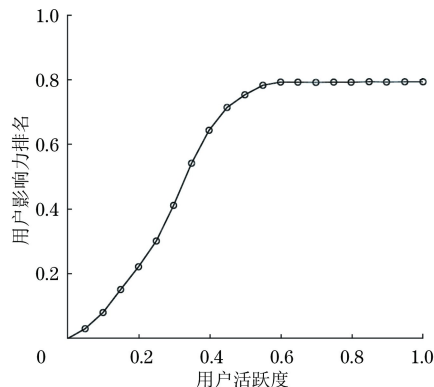


图 3 用户活跃度与用户影响力的关系

Fig.3 Relationship between user activity degree and user influence

本文对 5 组数据集进行分类测试,并对 NBC 算法进行了误差比较,结果如图 4 所示。从图 4 可以看出,本文算法的误差曲线始终在 0 线上方,这就说

明本文所提的算法的分类结果比 NBC 和 TwitterRank 算法更好,且本文所提算法的曲线落

差值小于 NBC 和 TwitterRank 算法,这也就说明本文算法更趋于稳定。

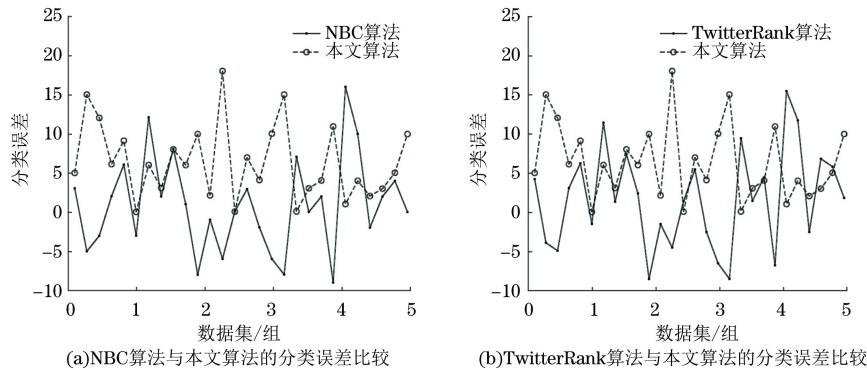


图 4 NBC、TwitterRank 算法和本文算法分类误差比较

Fig.4 Comparison of classification errors between NBC, TwitterRank algorithm and propose algorithm

最后对用户覆盖度的影响因素进行探讨,根据计算公式可以得出用户覆盖度和微博数呈现正相关。通过改变实验环境比较不同的活跃度对计算结果是否产生影响。从图 5 可知,活跃度越大,用户覆盖度越广,因此,用户的活跃度会对用户覆盖度产生影响。用户微博活跃粉丝数量会受到用户是否更新动态的影响,从侧面反映了用户的影响力。当微博数较少时,活跃度的影响力度比较明显,当微博数较多时,活跃度也会对覆盖度产生较大的影响力。

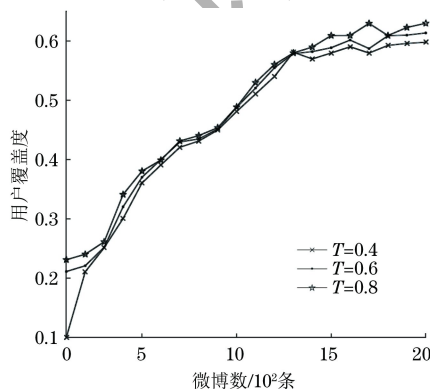


图 5 不同活跃度下用户覆盖度和微博数比较

Fig.5 Comparison of user coverage and Weibo number under different activity degree

4 结束语

针对社交网络用户影响力的评价问题,本文结合高斯贝叶斯衍生分类器提出了一种评价模型。该模型首先结合活跃度、联系度、覆盖度等指标设计社交网络用户影响力刻画模型,同时针对社交网络用户之间的关系特征和用户自身的行为特征,利用高斯贝叶斯衍生分类器给出上述模型的求解方法。最后以新浪社交网络用户作为实验数据,利用仿真实验研究了影响该评价模型的关键因素。实验结果表

明,相比 NBC 算法,该算法具有较好适应性。后续研究可以考虑利用社交网络信息的传播特征来完善用户影响力评价模型。

参考文献

- [1] CHA M, HADDADI H, BENEVENUTO F, et al. Measuring user influence in twitter: the million follower fallacy[C]//Proceedings of International AAAI Conference on Web and Social Media. Menlo Park, USA: AAAI Press, 2010: 10-17.
- [2] 徐文涛,刘锋,朱二周. 基于 MapReduce 的新型微博用户影响力排名算法研究[J]. 计算机科学, 2016, 43(9): 66-70, 86.
XU W T, LIU F, ZHU E Z. Research on novel ranking algorithm of microblog user's influence based on MapReduce [J]. Computer Science, 2016, 43(9): 66-70, 86. (in Chinese)
- [3] 吴少华,马晓娟,胡勇. 基于改进 PageRank 算法的微博用户影响力评估[J]. 四川大学学报(自然科学版), 2015, 52(5): 1040-1044.
WU S H, MA X J, HU Y. Impact evaluation of microbloggers based on improved PageRank algorithm[J]. Journal of Sichuan University (Natural Science Edition), 2015, 52(5): 1040-1044. (in Chinese)
- [4] 苑卫国,刘云. 微博用户特征量增长规律研究[J]. 计算机研究与发展, 2015, 52(2): 522-532.
YUAN W G, LIU Y. Growth law of user characteristics in microblog [J]. Journal of Computer Research and Development, 2015, 52(2): 522-532. (in Chinese)
- [5] 黄发良,冯时,王大玲,等. 基于多特征融合的微博主题情感挖掘[J]. 计算机学报, 2017, 40(4): 872-888.
HUANG F L, FENG S, WANG D L, et al. Mining topic sentiment in microblogging based on multi-feature fusion[J]. Chinese Journal of Computers, 2017, 40(4): 872-888. (in Chinese)
- [6] 刘冰玉,王翠荣,王聪,等. 基于动态主题模型融合多维数据的微博社区发现算法[J]. 软件学报, 2017, 28(2): 246-261.
LIU B Y, WANG C R, WANG C, et al. Microblog community discovery algorithm based on dynamic topic model with multidimensional data fusion[J]. Journal of Software, 2017, 28(2): 246-261. (in Chinese)
- [7] 张仰森,郑佳,唐安杰. 基于多特征融合的微博用户权威度定量评价方法[J]. 电子学报, 2017, 45(11): 2800-2809.
ZHANG Y S, ZHENG J, TANG A J. A quantitative evaluation method of micro-blog user authority based on

- multi-feature fusion[J]. Acta Electronica Sinica, 2017, 45(11): 2800-2809. (in Chinese)
- [8] 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 X, PENG X Q, HU M, et al. Extended multi-modality features and deep learning based microblog short text sentiment analysis[J]. Journal of Electronics & Information Technology, 2017, 39(9): 2048-2055. (in Chinese)
- [9] 李洋,陈毅恒,刘挺. 微博信息传播预测研究综述[J]. 软件学报, 2016, 27(2): 247-263.
LI Y, CHEN Y H, LIU T. Survey on predicting information propagation in microblogs[J]. Journal of Software, 2016, 27(2): 247-263. (in Chinese)
- [10] 王双成,杜瑞杰,刘颖. 连续属性完全贝叶斯分类器的学习与优化[J]. 计算机学报, 2012, 35(10): 2129-2138.
WANG S C, DU R J, LIU Y. The learning and optimization of full Bayes classifiers with continuous attributes [J]. Chinese Journal of Computers, 2012, 35(10): 2129-2138. (in Chinese)
- [11] 房丙午,黄志球,李勇,等. 基于贝叶斯网络的复杂系统动态故障树定量分析方法[J]. 电子学报, 2016, 44(5): 1234-1239.
FANG B W, HUANG Z Q, LI Y, et al. Quantitative analysis method of dynamic fault tree of complex system using Bayesian network[J]. Acta Electronica Sinica, 2016, 44(5): 1234-1239. (in Chinese)
- [12] 程玉虎,全瑶瑶,王雪松. 类相关性影响可变选择性贝叶斯分类器[J]. 电子学报, 2011, 39(7): 1628-1633.
CHENG Y H, TONG Y Y, WANG X S. A selective Bayesian classifier based on change of class relevance influence[J]. Acta Electronica Sinica, 2011, 39(7): 1628-1633. (in Chinese)
- [13] 王双成,高瑞,杜瑞杰. 小时间序列的动态朴素贝叶斯分类器学习与优化[J]. 控制与决策, 2017, 32(1): 163-166.
WANG S C, GAO R, DU R J. Learning and optimization of dynamic naive Bayesian classifiers for small time series[J]. Control and Decision, 2017, 32(1): 163-166. (in Chinese)
- [14] 易明,赵锦香,张展豪,等. 人类动力学视角下在线社区信息交流行为研究[J]. 情报科学, 2020, 38(8): 3-10.
YI M, ZHAO J X, ZHANG Z H, et al. Research on information communication behavior in online community based on human dynamics[J]. Information Science, 2020, 38(8): 3-10. (in Chinese)
- [15] 王亚奇,王静,杨海滨. 基于复杂网络理论的微博用户关系网络演化模型研究[J]. 物理学报, 2014, 63(20): 408-414.
WANG Y Q, WANG J, YANG H B. An evolution model of microblog user relationship networks based on complex network theory[J]. Acta Physica Sinica, 2014, 63(20): 408-414. (in Chinese)
- [16] 龙增艳,陈志刚,徐成林. 基于用户交互的社交网络好友推荐算法[J]. 计算机工程, 2019, 45(3): 132-137.
LONG Z Y, CHEN Z G, XU C L. Social network friend recommendation algorithm based on user interaction [J]. Computer Engineering, 2019, 45(3): 132-137. (in Chinese)
- [17] 阳雨,胡亚洲,郭勇,等. 基于在线社交网络的用户信任传递建模与分析[J]. 计算机工程, 2018, 44(11): 265-270.
YANG Y, HU Y Z, GUO Y, et al. Modeling and analysis of user trust transfer based on online social network [J]. Computer Engineering, 2018, 44(11): 265-270. (in Chinese)
- [18] MATSUMURA N, OHSAWA Y, ISHIZUKA M. Influence diffusion model in text-based communication [J]. Transactions of the Japanese Society for Artificial Intelligence, 2002, 17: 259-267.
- [19] 郭宏刚,杨芳. 基于影响力的跨社交网络谣言扩散模型与抑制方法[J]. 计算机应用与软件, 2022, 39(7): 73-79, 153.
GUO H G, YANG F. Rumor spreading model and reduce method for cross online social networks based on influence [J]. Computer Applications and Software, 2022, 39(7): 73-79, 153. (in Chinese)
- [20] 毛佳昕,刘奕群,张敏,等. 基于用户行为的微博用户社会影响力分析[J]. 计算机学报, 2014, 37(4): 791-800.
MAO J X, LIU Y Q, ZHANG M, et al. Social influence analysis for micro-blog user based on user behavior [J]. Chinese Journal of Computers, 2014, 37(4): 791-800. (in Chinese)
- [21] 石立鹏,王莉. 基于遍历约束与交互信息增强的社交网络表征算法[J]. 计算机工程, 2018, 44(11): 215-221.
SHI L P, WANG L. An enhanced social network representation algorithm based on traversal constraint and interactive information[J]. Computer Engineering, 2018, 44(11): 215-221. (in Chinese)
- [22] XIAO F, NORO T, TOKUDA T. Finding news-topic oriented influential twitter users based on topic related hashtag community detection [J]. Journal of Web Engineering, 2014, 13(5/6): 405-429.
- [23] NORO T, TOKUDA T. Searching for relevant tweets based on topic-related user activities [J]. Journal of Web Engineering, 2016, 15(3/4): 249-276.
- [24] 谢柏林,黎琦,邝建. 基于隐半马尔可夫模型的微博流行信息检测方法[J]. 计算机科学, 2022, 49(S1): 291-296.
XIE B L, LI Q, KUANG J. Detection method of popular information in weibo based on hsmm[J]. Computer Science, 2022, 49(S1): 291-296. (in Chinese)
- [25] 徐志明,李栋,刘挺,等. 微博用户的相似性度量及其应用[J]. 计算机学报, 2014, 37(1): 207-218.
XU Z M, LI D, LIU T, et al. Measuring similarity between microblog users and its application[J]. Chinese Journal of Computers, 2014, 37(1): 207-218. (in Chinese)
- [26] 崔增乐,钱晓东. 区块链社交网络信息传播模型的优化研究[J]. 计算机工程与应用, 2021, 57(7): 59-69.
CUI Z L, QIAN X D. Research on optimization of information spreading model of blockchain social network[J]. Computer Engineering and Applications, 2021, 57(7): 59-69. (in Chinese)

编辑 索书志