

面向推荐系统的用户评论体验提取方法

粟勇^{1,2}, 谢裕清³, 孙华成⁴, 王远⁴, 王永利⁴

1. 南瑞集团有限公司(国网电力科学研究院有限公司), 江苏南京 211100;
2. 江苏瑞中数据股份有限公司, 江苏南京 211100;
3. 国网浙江省电力有限公司信息通信分公司, 浙江杭州 310000;
4. 南京理工大学计算机科学与工程学院, 江苏南京 210094)

摘要:推荐技术能够根据用户的历史行为等客观数据推测用户的偏好,为电商营销和运维提供决策依据和指导。为了提高特定电子商务领域推荐的准确度,从不同角度研究分析了用户评论体验,加强对用户语言模式的本体研究,有效区分主题和非主题产品特征,并给出用户体验的符号描述。提出了一种面向推荐系统的用户体验提取算法,将用户体验特征词、副词、修饰语构建的短语作为具有推荐解释功能的标签,弥补了传统的标签推荐方法需要依赖标签数据源的问题。同时提出用户体验概念,给出提取算法流程,设计了针对评论内容的协同滤波算法,可以提取用户体验情感极性用于用户评分,结合情感极性与体验特征还可以应用于基于张量的推荐系统。实验结果表明,提取的用户体验同时具有准确性和多样性,并且能够保证较高的准确率与查全率。

关键词:推荐系统;偏好;用户评论;用户体验;协同滤波;情感极性

中图分类号:TP391.1

文献标识码:A

文章编号:1673-629X(2022)06-0052-05

doi:10.3969/j.issn.1673-629X.2022.06.009

User Comment Experience Extraction Method for Recommendation System

SU Yong^{1,2}, XIE Yu-qing³, SUN Hua-cheng⁴, WANG Yuan⁴, WANG Yong-li⁴

1. NARI Group Corporation (State Grid Electric Power Research Institute), Nanjing 211100, China;
2. China Realtime Database Co., Ltd., Nanjing 211100, China;
3. Information & Telecommunication Branch, State Grid Zhejiang Electric Power Corporation, Hangzhou 310000, China;
4. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Recommendation technology can predict users' preferences based on objective data such as their historical behaviors, providing decision-making basis and guidance for e-commerce marketing, operation and maintenance. In order to improve the accuracy of recommendation in specific e-commerce field, we study and analyze the user comments experience from different perspectives, strengthen the ontology research of user language patterns, effectively distinguish the features of theme and non-theme products, and give the symbolic description of user experience. We propose a user experience extraction algorithm for recommendation system. The phrases constructed by user experience feature words, adverbs and modifiers are used as tags with recommendation interpretation function for tag recommender system, which makes up for the problem that traditional tag recommender system relies on tag data source. Furthermore, the concept of user experience is proposed, the extraction algorithm flow is given, and a collaborative filtering algorithm for comment content is designed, which can extract the emotional polarity of user experience for user rating. Combined with the information of emotional polarity and experience characteristics, it can also be used in tensor-based recommendation system. Experimental results show that the extracted user experience has both precision and diversity, and can ensure high precision and recall.

Key words: recommendation system; preferences; user comments; user experience; collaborative filtering; emotional polarity

0 引言

随着大数据的快速发展,推荐系统已经被成功应

用到电商大数据领域,用以挖掘用户的历史行为、用户生成内容,并在此基础上建模来推测用户偏好^[1-3],并

收稿日期:2021-03-22

修回日期:2021-07-23

基金项目:国家电网科技项目(5211XT190033)

作者简介:粟勇(1985-),男,硕士研究生,研究方向为电力信息化。

向其推荐可能符合需求的信息。在推荐系统中,用户的主观评价和商品的特征信息并不是以一种约定俗成的形式或者结构存在于评论之中,更多的是以非结构化形式存在,在构建用户画像并将其运用于推荐系统这一方面,对于如何挖掘这些非结构化信息中实际可用的价值,很多学者提出了很多方法。例如,Sheng Feng 等^[4]提出一种选取合适的用户评论用于产品推荐的方法,但是该方法并没有区分主题产品和非主题产品各自的特征。Valliyammai C^[5]提出了一种基于用户评论的旅游推荐系统,根据其他用户的历史评价推荐最佳的旅游方案,但这种方式所采用的支持向量机(support vector machine, SVM)占用了大量内存并且其个性化程度高度依赖于同义词库。Chakraverty 等^[6]提出了一种跨领域推荐方法。当一个领域中的可用数据缺少时,可以使用来自其他领域的数据(用户评分和评价)进行跨领域推荐。

为了有效区分主题产品特征和非主题产品特征,加强对用户语言模式的本体研究,在现有研究^[7-14]的基础上,该文将非结构化的评论信息提取成结构化的多维用户体验信息,提出了一种在评论中提取用户体验的方法。详细介绍了该方法具体的步骤和算法,有利于推荐系统提供更准确的个性化推荐结果。实验结果表明该方法能够实现较高的准确率和丰富的多样性。

1 面向推荐系统的用户体验提取算法

1.1 评论中的用户体验

文献[15]提出了一种能够提取用户体验的自动化方法(AEUE)。其中,该方法定义的用户体验是由提取并过滤后的产品特征词和特征描述词组成。进一步给出用户体验的定义如下:

定义1 用户体验(E):用户体验由一个五元组表示: \langle 体验特征(E_n),程度副词(E_{adv}),修饰语(E_a),情感极性(E_v),权重(E_w) \rangle ,用该五元组标签表示用户的体验信息,例如“价格很实惠”、“实用性很强”等。 E_n 表示名词,如上述中的“价格”、“实用性”, E_a 表示修饰词,如“实惠”、“强”等。 E_{adv} 表示修饰 E_n 的副词,如“很”、“较”。 E_v 为情感极性的数值。

表2 句子依存关系分类及解释

关系类型	标签	描述	示例
间宾关系	IOB	indirect-object	我教他英语(教->英语)
动宾关系	VOB	verb-object	我教他英语(教->他)
主谓关系	SBV	subject-verb	我喜欢小动物(我<-喜欢)
兼语	DBL	double	我命令你离开(命令->你)
前置宾语	FOB	fronting-object	饭吃完了(饭<-吃)

E_w 表示用户体验权重。

1.2 基于评论的用户体验提取模型

基于评论的挖掘的目的是在非结构性的用户评价语中通过将产品的特征提取出来,构建成特征向量。该文在实际处理过程中将产品特征划分为功能性特征和描述性特征两个方面,下面对这两类特征做简单的介绍:

(1)功能性特征:产品本体元数据,如“电量持久”、“前置摄像清晰”、“运行内存大”等就是平板电脑的功能性特征描述。

(2)描述性特征:对功能特征的描述,如“8 000 mA 电池容量”、“2 500 万前置摄像”、“8 G 运行内存”等。

受 Schein 等人^[14]的启发,该文将用户体验提取过程总结为如下几个主要步骤:

- (1)构建种子词集;
- (2)提取体验特征;
- (3)匹配出产品特征的修饰性短语或语段;
- (4)采用一定方法分析出(3)中短语或语段的情感倾向,主要分为褒义和贬义;
- (5)计算用户体验的权重并筛选排序。

下面对上述步骤和其中牵涉的主要算法进行详细的介绍。

1.2.1 构建种子词集

在预处理过程中,分词之后根据词性标注和词频统计排序,再人工选出相关度高的词,由此产生种子。表1列举了部分种子词。

表1 种子词集合

用户分类	种子词
年龄段	60后、70后、90后等
地域	华南、华北、华中等
生命周期	考察期、成长期、成熟期、衰老期
访问频率	高频访问、中频访问、低频访问

1.2.2 提取体验特征

短语结构文法主要是完成对句子各成分结构化关系的构建。依存文法则是通过句子中词语之间的依存关系来描述句子的结构,主要是中心词和其他句子成分之间的关系。依存文法中具体包含的关系种类由表2给出。

续表 2

关系类型	标签	描述	示例
定中关系	ATT	attribute	咸鸭蛋 (咸<-鸭蛋)
动补结构	CMP	complement	看过了电视 (看->过)
状中结构	ADV	adverbial	十分壮观 (十分<-壮观)
右附加关系	RAD	right adjunct	海浪和沙滩 (海浪->和)
左附加关系	LAD	left adjunct	海浪和沙滩 (和<-沙滩)
介宾关系	POB	preposition-object	为人民服务 (为->人民)
并列关系	COO	coordinate	你和我 (你->我)
核心关系	HED	head	整个句子的核心
独立结构	IS	independent structure	两个单句在结构上彼此独立

定义 2 体验特征标签 (T)。体验特征标签由一个三元组表示, 即: \langle 体验特征 (E_n), 程度副词 (E_{adv}), 修饰语 (E_a) \rangle , 这个三元组作为中间提取结果。

依存分析后, 挖掘每个产品的候选标签集合 T , 即有:

$$T = \{T_1, T_2, \dots, T_n\}$$

$$T_i = \{t_1, t_2, \dots, t_m\}$$

$$t_i = \langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle, i \in [1, m]$$

其中, n 为产品 p 评论数, T_i 为第 i 个评论的特征标签。此外还需对三元组做显式去重操作。

基于评论的特征提取方法由算法 1 描述。

算法 1: 基于评论的特征提取方法。

输入: 产品集合 S ; 产品 S 的评论集合 r_s ; 种子词集合 W

输出: $T: \langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle$

- 1: for s in S do
- 2: for r in r_s do
- 3: preprocess ()
- 4: DP \leftarrow DependencyParser(r_s) // 依存文法分析
- 5: if DP includes $\langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle$ and $n \in W$ and $a \in W$ then
- 6: // 获取 $\langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle$ 和包含种子词的 ADV, SBV 依存关系
- 7: $T_{pr} \leftarrow \langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle$
- 8: end if
- 9: end for
- 10: end for

算法 1 中预处理的算法复杂度为 $O(|S|)$, 依存文法分析的算法复杂度平均值一般为 $O(|r|^3)$, 对应本算法的整体复杂度为 $O(|S||R||\bar{r}|^3)$, 其中 $|S|$ 为整个语料的文字数目, $|\bar{r}|$ 为平均评论的句子长度, $|R|$ 为整体的评论数量。

1.2.3 情感极性计算

定义 3 情感极性 (E_v): 指根据 E_n 与 E_{adv} 的情感分析得到的、并经过归一化处理后的语义倾向。

E_v 计算方法由算法 2 给出:

算法 2: 情感极性计算方法。

输入: $T: \langle \text{word}_n, \text{word}_{adv}, \text{word}_a \rangle$; 情感词种子词集合 W ;

输出: E , 包括 $E_n, E_{adv}, E_a, E_v, E_w$ 的五元组。

- 1: for t in T do
- 2: calculate PMI($t, \text{word}_{positive}$)
- 3: calculate PMI($t, \text{word}_{negative}$)
- 4: $x \leftarrow SO(t)$ // 计算语义倾向
- 5: if $x \geq \text{threshold}_1$ and $x \leq \text{threshold}_2$ then // 根据阈值筛选体验标签
- 6: $E(w)_{pr} = \langle T_{pr}, x \rangle$ // 过滤体验特征标签
- 7: end if
- 8: end for
- 9: Normalization(E)

算法 2 整体的算法复杂度为 $O(|W||T|)$, $|T|$ 为用户体验特征标签的数目, $|W|$ 为种子词的数目。

1.2.4 用户体验聚类方法

该文主要采用基于隐含狄利克雷分布 (latent Dirichlet allocation, LDA) 的主题方式进行聚类。使用聚类的目的是让提取的用户体验具有多样性, 同时聚类之后还要确保类簇之间保持独立性。

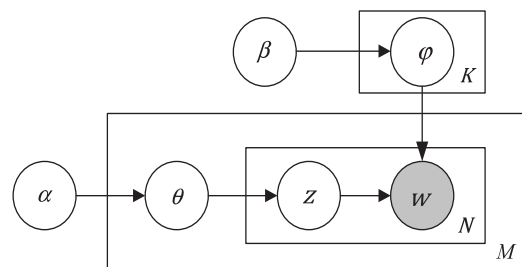


图 1 LDA 的概率图模型

如图 1 中表述的是 LDA 的概率图模型, 其中 N 表示训练语料库中出现的所有词, Z 表示隐藏的主题, M 表示文档数目, W 表示可观测词, K 表示主题个数, α, β 分别表示每篇文档中主题和每个主题的先验狄利克雷分布的超参数, $\theta \in R^{M \times K}$ 表示第 m 篇文章中的主题分布, φ 表示第 k 个主题中的主题词分布。

由图可知,LDA 联合概率为:

$$p(\theta, z, w | \alpha, \beta) = p(\theta | \alpha) \prod_{n=1}^N p(z_n | \theta) p(w_n | z_n, \beta) \quad (1)$$

算法 3 描述了主题聚类算法详细过程。

算法 3:LDA 主题聚类算法。

输入:主题数目 L ;文章集合 D ,即标签词;超参数 α, β ;

输出:每隔主题下的词概率分布 φ ;每篇文章的主题概率分布 θ ;每个主题下 φ 概率从高到低 topN 个特征词 W_{topN}

```

1: for topics  $l \in [1, L]$  do
2:   sample mixture components  $\vec{\varphi} \sim \text{Dir}(\vec{\beta})$ 
3: end for
4: for documents  $m \in [1, M]$  do
5:   sample mixture proportion  $\vec{\vartheta}_m \sim \text{Dir}(\vec{\alpha})$ 
6:   sample document length  $N_m \sim \text{Poiss}(\xi)$ 
7:   for words  $n \in [1, N_m]$  do
8:     sample topic index  $Z_{m,n} \sim \text{Mult}(\vec{\vartheta}_m)$ 
9:     sample term for word  $W_{m,n} \sim \text{Mult}(\vec{\varphi}_{z_{m,n}})$ 
10:  end for
11: end for
    
```

1.2.5 用户体验权重计算方法

本节介绍权重的计算和处理。经过上述聚类后就可以得到类簇,再根据类簇的权重大小提取用户体验。最后将归一化权重值赋给 E_w 。权重计算公式如下:

$$E_{w_i} = \frac{|z_i|}{|T|} \quad (2)$$

其中, $|T|$ 为产品中体验特征标签总数, z_i 为 t_i 标签所属的主题, $|z_i|$ 为主题中主题词的数目。

2 实验与结果分析

2.1 实验环境

实验在 python2.7 环境下完成。分词使用的是 jieba 分词,版本号为 0.39。依存文法分析使用的是哈工大语言技术平台的开发接口,采用 gensim3.7.3 作为 LDA 主题聚类模型库。

2.2 实验数据

从淘宝的天猫商城选取 50 个不同种类的产品数据作为实验数据,该文展示的主要是八种部分商品,包括“iPhone X”、“OPPO R11S”、“拜耳电动牙刷”、“吉利剃须刀”、“零食礼包”等等。

2.3 评价指标

为了验证提出方法的有效性,采用准确率 (precision) 和召回率 (recall) 来评估体验提取的效果。对于产品 $i \in I$, $TP(i)$ 表示该算法对于产品评论集合中提取出的用户体验符合产品特征的集合, $FP(i)$ 表示提取出的用户体验与产品特征不相关的集合。推荐结果的 precision 和 recall 的定义分别为:

$$\text{precision} = \frac{\sum_{i \in I} |TP(i) \cap FP(i)|}{\sum_{i \in I} |TP(i)|} \quad (3)$$

$$\text{recall} = \frac{\sum_{i \in I} |TP(i) \cap FP(i)|}{\sum_{i \in I} |FP(i)|} \quad (4)$$

并采用 F1 来评价分类模型的好坏:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5)$$

该文采用修正召回率,计算方式如表 3 所示。

表 3 查重率计算

数据来源	总数	查准率	查重率	修正召回率
result ₁	A ₁	C ₁	$\frac{C_1}{A_1}$	$\frac{C_1}{C_1 + C_2}$
result ₂	A ₂	C ₂	$\frac{C_2}{A_2}$	$\frac{C_2}{C_1 + C_2}$

将修正召回率计算得到的 F1 值称为修正的 F1 值,记为 F_w ,传统的 F1 值记为 F_l 。当满足 $C_1 \geq C_2$ 时:

$$F_w(\text{result}_1) > F_w(\text{result}_2) \Rightarrow F_l(\text{result}_1) > F_l(\text{result}_2) \quad (6)$$

在实验过程中, result₁ 作为提取结果, result₂ 作为对比方法提取的结果。多样性使用语义多样性计算:

$$\text{Diversity} = \frac{N_{\text{indep}_s}}{N_{\text{right}_s}} \quad (7)$$

其中, N_{indep_s} 表示用户体验的实际标签数, N_{right_s} 表示用户体验的正确标签数。

表 4 体验信息特征标签提取结果

产品	实际标签数	正确数	评论数量
iPhone X 手机	5	5	12 658
欧普(OPPO) R11S	5	4	3 517
拜耳电动牙刷	5	5	19 460
吉利剃须刀	5	4	13 847
零食礼包	5	4	77 933
senma 卫衣	5	5	3 558
耐克(空军一号)	5	5	1 365
欧式沙发	5	5	1 839

2.4 结果分析

为了分别验证体验信息提取的准确性和多样性,设计了两组实验。第一组实验的标签提取结果如表 4 所示。每类产品选取前 5 个最相关的体验信息标签,比如对于 iPhone X 的 5 个标签为“正品行货”、“操作方便”、“性能强劲”、“屏幕清晰”、“款式漂亮”。实验结果表明文中方法的平均准确率为 92.5%。

第二组实验的实验结果由表 5 给出。表 5 中方法 1 代表文中提出的基于推荐系统的用户评论体验提取方法,方法 2 表示文献[15]中用户体验信息自动提取

的方法。可以看到,文中方法最终准确率平均值为 0.872 5,略高于 AEUE 方法的准确率(0.832 5),但是修正召回率和修正 F-measure 值低于 AEUE。所以

AEUE 在修正召回率计算上更具优势。而采用的 LDA 聚类保证类簇之间的独立性,故文中方法在多样性上更具优势。

表 5 用户体验提取结果对比

产品名称	提取方法	体验信息总数	正确数量	多样性	修正召回率	修正 F-Measure	准确率
iPhone X 手机	1	37	33	0.58	0.45	0.59	0.89
	2	54	41	0.54	0.55	0.64	0.76
欧普(OPPO)R11S	1	41	36	0.58	0.43	0.58	0.88
	2	58	47	0.70	0.57	0.67	0.81
拜耳电动牙刷	1	27	23	0.83	0.35	0.49	0.85
	2	45	43	0.53	0.65	0.77	0.96
吉利剃须刀	1	19	14	0.57	0.19	0.31	0.74
	2	63	58	0.21	0.81	0.86	0.92
零食礼包	1	22	20	0.65	0.38	0.53	0.91
	2	45	33	0.58	0.62	0.67	0.73
senma 卫衣	1	21	18	0.83	0.29	0.43	0.86
	2	48	44	0.64	0.71	0.80	0.92
耐克(空军一号)	1	29	28	0.68	0.36	0.52	0.97
	2	53	50	0.52	0.64	0.76	0.94
欧式沙发	1	24	21	0.86	0.47	0.61	0.88
	2	39	24	0.79	0.53	0.57	0.62

3 结束语

为了提高推荐准确度,改善推荐质量,提出一种基于评论的用户体验提取算法。实验结果证明了该算法的有效性,并且能够实现较高的准确率与查全率。

参考文献:

- [1] 马 亮,陶利涛,谢骏凯. 基于客户画像的客户诉求管理[J]. 电力需求侧管理,2016,18(21):98-100.
- [2] SORBO A D, PANICHELLA S, ALEXANDRU C V, et al. SURF: summarizer of user reviews feedback [C]//39th IEEE international conference on software engineering (ICSE 2017). Buenos Aires, Argentina; IEEE, 2017: 55-58.
- [3] WANG F, CHEN L. Recommending inexperienced products via learning from consumer reviews [C]//IEEE/WIC/ACM international conferences on web intelligence & intelligent agent technology. Macau, China; ACM, 2012: 596-603.
- [4] FENG S, ZHANG M, ZHANG Y, et al. Recommended or not recommended? review classification through opinion extraction [C]//12th international Asia-Pacific web conference. Busan, Korea (South); IEEE, 2010: 350-352.
- [5] VALLIYAMMAI C, PRASANNAVENKATESH R, VENNILA C, et al. An intelligent personalized recommendation for travel group planning based on reviews [C]//2016 eighth international conference on advanced computing (ICoAC). Chennai, India; IEEE, 2017: 67-71.
- [6] CHAKRAVERTY S. On using reviews and comments for cross domain recommendations and decision making [C]//3rd international conference on computing for sustainable

global development (Indiacomm 2016). Mumbai, India; IEEE, 2016: 3656-3659.

- [7] 姜书浩,张立毅,周 娜. 个性化推荐系统中的多样性综述[J]. 软件工程与应用,2019,8(3):172-178.
- [8] 肖 倩,韩 婷,张 聪. 社会化媒体环境中的数字阅读物推荐及其用户体验研究—以豆瓣阅读为例[J]. 科技与出版, 2014(11):65-69.
- [9] ZHANG X, DING D, HUANG C. A user intention modeling algorithm for friend recommendation [C]//2017 IEEE 2nd international conference on big data analysis (ICBDA). Beijing, China; IEEE, 2017.
- [10] GUO Dongyan, XU Jingsong, ZHANG Jian, et al. User relationship strength modeling for friend recommendation on Instagram [J]. Neurocomputing, 2017, 239: 9-18.
- [11] HAN Jiayu, ZHENG Lei, XU Yuanbo, et al. Adaptive deep modeling of users and items using side information for recommendation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(3): 737-748.
- [12] CHU Pang-Ming, MAO Yu-Shun, LEE Shie-Jue, et al. Leveraging user comments for recommendation in e-commerce [J]. Applied Sciences, 2020, 10(7): 2540.
- [13] SCHEIN A I, POPESCU A, UNGAR L H, et al. Methods and metrics for cold-start recommendations [C]//Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval (SIGIR 2002). Tampere, Finland; ACM SIGIR, 2002: 253-260.
- [14] 李 阜. 基于滑窗取词的单文档自动摘要技术研究 [D]. 长沙:国防科学技术大学, 2010.
- [15] 胡令传,陶晓鹏. 客户评论中用户体验信息自动提取研究 [J]. 计算机工程, 2015, 41(1): 49-53.