

基于词向量的产品评论有用度评估方法

郑华飞, 周向东

(复旦大学 计算机科学技术学院, 上海 200433)

摘 要: 产品评论的快速增长以及质量的参差不齐, 使得消费者获得有用的产品评论变得困难。为此, 提出一种新的产品评论有用度评估方法。引入词向量作为评论文本的深度特征表示, 结合结构特征、情感特征、元数据特征等训练回归模型, 自动地对评论进行有用度评估并基于有用度对评论进行排序。在 Amazon 真实数据集上的实验结果表明, 该方法在回归性能和排序性能上均优于 UGR + LEN + STR 方法和基准方法。另外通过挖掘特定领域的词向量特征, 该向量模型在 RMSE, NDCG 等评价指标上可有效地改善评估效果。

关键词: 产品评论; 词向量; 有用度; 深度学习; 神经网络语言模型

中文引用格式: 郑华飞, 周向东. 基于词向量的产品评论有用度评估方法[J]. 计算机工程, 2017, 43(4): 188-193.

英文引用格式: Zheng Huafei, Zhou Xiangdong. Product Review Helpfulness Assessment Method Based on Word Vector[J]. Computer Engineering, 2017, 43(4): 188-193.

Product Review Helpfulness Assessment Method Based on Word Vector

ZHENG Huafei, ZHOU Xiangdong

(School of Computer Science, Fudan University, Shanghai 200433, China)

【Abstract】 Rapid increasing volume of product reviews as well as their variant qualities makes it time-consuming and tedious for individuals to perceive valuable reviews. Therefore, this paper proposes an approach to automatically assess helpfulness of online product reviews based on word vector. The approach introduces word vector as a deep text feature and incorporates it with structure feature, sentiment feature and meta feature to learn a regression model for automatically helpfulness assessing. A ranking procedure based on helpfulness is performed. Compared with UGR + LEN + STR's approach and the baseline, experimental results conducted on Amazons dataset show that this approach achieves promising performances both on regression and ranking. Furthermore, this paper explores domain-specific word vector model, which can improve assessing effect on RMSE, NDCG and other evaluation indexes.

【Key words】 product reviews; word vector; helpfulness; deep learning; neural network language model

DOI: 10.3969/j.issn.1000-3428.2017.04.032

0 概述

近年来电子商务的发展非常活跃, 越来越多的用户倾向于在网上购买商品和服务并进行评论。产品评论数据蕴含了非常有价值的商业信息, 相比于商家提供的图片和口号, 用户更愿意相信有过购买和使用经验的用户所发表的观点。互联网是一个自由开放的环境, 用户可以提交任何评论, 而随着评论数据的日益增长, 内容虚假、包含广告、甚至是无意义的随机文本等垃圾评论^[1]严重影响了产品评论的整体质量。此外, 受欢迎的产品往往拥有非常多的评论, 用户如果对每一条评论进行阅读将花费大量

的时间和精力。因此, 为了让用户更有效地感知产品评论, 需要开展产品评论的有用度评估技术研究。尽管某些网站, 如 Amazon.com 提供了用户对评论有用度的投票 (vote) 机制, 但产品评论的有用度评估依然非常必要: 1) 有用度评估技术可以感知在根本上影响有用度的因素, 避免 Early Bird 偏差^[2]。2) 不具备投票机制的电商网站可以通过有用度评估技术挖掘用户评论, 改善用户体验。

近年来, 深度学习在自然语言处理中的应用越来越广泛, 利用神经网络训练的词向量模型被证明可以提高如词性标注、命名实体识别、语义角色标注等传统自然语言处理任务的效果^[3], 这给产品评论

基金项目: 国家自然科学基金 (61370157); 上海市科技计划项目 (14511107403)。

作者简介: 郑华飞 (1990—), 男, 硕士研究生, 主研方向为自然语言处理、信息检索、文本挖掘; 周向东, 教授。

收稿日期: 2016-03-15 **修回日期:** 2016-05-03 **E-mail:** vincentzhengcs@126.com

相关的诸多应用带来了新的机遇和挑战。本文引入深度学习技术,提出一个新的基于词向量的产品评论有用度评估方法,利用词向量所携带的更丰富的语义信息挖掘更深层次的评论文本特征表示。

1 相关工作

在产品评论有用度评估问题上,前人的工作主要侧重于多方面地挖掘产品评论的特征,诸如文本特征、书写风格、产品属性、用户偏好、相关情感等。文献[4]基于用户行为对垃圾评论进行检测,从垃圾评论者的目的出发,对 5 种垃圾评论行为模式进行线性组合计算垃圾评论者得分,然而该文献没有挖掘垃圾评论的语义特征。文献[5]利用 Amazon.com 产品评论数据,训练支持向量回归模型(SVR)来学习有用度得分函数并应用到排序系统中,并提出评论的长度、一元模型(unigram)和产品得分(star)等特征的有效性,但是该文献没有挖掘句法特征以及情感特征对于有用度评估的影响。文献[6]侧重于从产品评论中抽取用户体验信息,利用语法规则和共现概率分别抽取产品特征词和特征描述词,然而基于规则的特征抽取不具有迁移性,并且相对于数据的变化具有滞后性。文献[7]着重考察了可读性这一特征,并使用有监督的方法证明了可读性对于预测评论有用度的有效性。文献[8]对用户专长、写作风格、发布时间等做了详尽的分析,并提出了一种非线性的基于径向基的函数模型来预测有用度。文献[9]侧重于基于用户偏好的特征来改善前人基于文本特征的有效度评估模型,提出 3 种用户偏好特征:1)该评论是否满足用户对于产品属性(attributes)和功能(functions)的信息需求。2)评论所含信息的可靠性,通过能愿动词的不确定性和时态来衡量。3)主流观点倾向。但是文献[8-9]没有针对评论内容的语义进行理解。文献[10]研究了跨领域产品评论情感倾向分析问题,通过领域独立词建立起目标领域词和原领域词的关系,对评论文本进行特征变换,取得了较好的分类效果。文献[11]在文献[5]的基础上加入了情感特征,主要利用 LIWC 和 INQUIRER 两组情感映射词典将每条评论表示为直方图(histogram)的特征形式。

通过以上分析,从评论内容的语义角度出发对文本语义进行建模的工作尚不多见。本文利用神经网络语言模型,直接对评论文档语义进行建模,并通过基于向量空间的特征变换,抽取基于词向量的评论文本深度特征表示,在文献[5,11]的基础上,结合其他经典的一元模型特征 UGR、产品得分特征 STR、评论长度特征 LEN、情感特征 SENTI 等,提高有用度评估模型的效果。

2 产品评论有用度评估

2.1 问题定义

产品评论有用度评估任务旨在首先自动地对评论进行有用度投票,评估其有用度;然后根据所评估的有用度得分,对于给定的产品进行评论排序。本文参考文献[5]的方法定义产品评论的有用度数值:

$$h(r \in R) = \frac{rating_+(r)}{rating_+(r) + rating_-(r)} \quad (1)$$

其中, R 是评论文档集合; $rating_+(r)$ 表示给评论 r 投“有用”票的人数; $rating_-(r)$ 表示给其投“无用”票的人数,下面一条 Amazon 评论数据为例:

No carrier can active this phone, this phone was reported lost and stolen. (21/22)

其中, $rating_+(r)$ 为 21, $rating_-(r)$ 为 1, 表示 22 个评论浏览者中有 21 个人认为该条评论“有用”。本文使用“有用度”数值 $h(r \in R)$ 作为数据标签,用来训练回归模型。

2.2 文本深度特征

本文对基于词向量模型的文本深度特征在评论有用度评估中的作用进行研究与探讨。词向量由神经网络语言模型训练得到,与传统的统计语言模型(n-gram)相比,主要的优势有以下 2 点:

1) 词向量携带了更多的语法和语义信息,词与词之间的语法和语义上的相似性,可以通过词向量来体现。

2) 基于词向量的模型自带平滑功能。

2.2.1 词向量训练

本文使用基于 Negative Sampling 的 Continuous Bag-of-words 模型(Neg-CBOW)^[12-13]训练词向量。图 1 给出了 Neg-CBOW 模型的网络结构,包含三层:输入层,投影层和输出层,下面以样本($Context(w), w$)为例说明该网络的训练过程,其中向量 w 为当前词,向量 $Context(w)$ 为 w 的上下文,这里取 w 前后 c 个词,沿用前人的经验^[12],在实验中对于超参数 c 取 5。

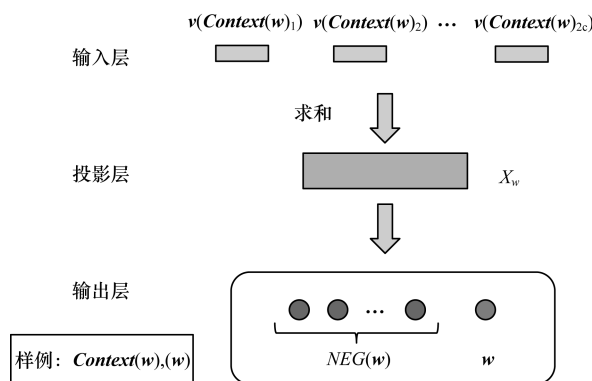


图 1 基于 Negative Sampling 的 CBOW 模型网络结构

1) 输入层

包含 $2c$ 个词的词向量,分别用 $v(Context(w)_1),$

$v(\text{Context}(\mathbf{w})_2), \dots, v(\text{Context}(\mathbf{w})_{2c}) \in \mathbb{R}^m$ 表示, 其中, m 为词向量的长度。在训练初始, 输入层的词向量均为随机初始化。

2) 投影层

该层输入的 $2c$ 个词向量作累加求和, 即:

$$\mathbf{x}_w = \sum_{i=1}^{2c} v(\text{Context}(\mathbf{w})_i) \in \mathbb{R}^m \quad (2)$$

3) 输出层

该层通过随机负采样选出正样本词 \mathbf{w} 对应的负样本词集合 $NEG(\mathbf{w}) \neq \emptyset$, 在训练过程中最大化一系列条件概率的乘积:

$$g(\mathbf{w}) = \prod_{u \in |\mathbf{w}| \cup NEG(\mathbf{w})} p(u | \text{Context}(\mathbf{w})) \quad (3)$$

其中:

$$p(\mathbf{w} | \text{Context}(\mathbf{w})) = \begin{cases} \sigma(\mathbf{x}_w^T \boldsymbol{\theta}^u), & u = \mathbf{w} \\ 1 - \sigma(\mathbf{x}_w^T \boldsymbol{\theta}^u), & \text{otherwise} \end{cases} \quad (4)$$

其中, $\boldsymbol{\theta}^u \in \mathbb{R}^m$ 表示词 u 对应的一个辅助参数向量。这里源于逻辑回归分类模型的思想, 定义一个节点被分为正类的概率是:

$$\sigma(\mathbf{x}_w^T \boldsymbol{\theta}^u) = \frac{1}{1 + e^{-\mathbf{x}_w^T \boldsymbol{\theta}^u}} \quad (5)$$

对于一个给定的样本 $(\text{Context}(\mathbf{w}), \mathbf{w})$, 最大化 $g(\mathbf{w})$ 相当于增大正样本的概率同时降低负样本的概率。对于给定的语料库 C , 将函数:

$$G = \prod_{\mathbf{w} \in C} g(\mathbf{w}) \quad (6)$$

作为 Neg-CBOW 神经网络整体优化目标, 为了计算方便取 $\log G$ 作为最终优化目标, 并使用随机梯度上升法进行优化求解。每取一个样本 $(\text{Context}(\mathbf{w}), \mathbf{w})$ 需要对目标函数中所有相关的参数即 $\mathbf{x}_w, \boldsymbol{\theta}^u, \mathbf{w} \in C$ 进行一次刷新。由于投影层为 $2c$ 个词向量作累加求和, 因此 \mathbf{x}_w 参数的更新需要贡献到输入层中每一个 $v(\text{Context}(\mathbf{w})_i)$ 上。

2.2.2 基于向量空间模型的特征空间变换

词向量是词的特征表示, 需要对其进行特征空间变换以表示文档特征。本文采用一种基于向量空间模型 (Vector Space Model, VSM) 的特征空间变换方法, 得到表示文档的词向量特征。

对于给定文档 d , 其基于统计的向量空间模型 (Statistical-VSM) 的特征向量表示为:

$$\mathbf{x}^{\text{Stat}}(d) = (x_1^{\text{Stat}}(d), x_2^{\text{Stat}}(d), \dots, x_M^{\text{Stat}}(d)) \quad (7)$$

其中, M 为向量空间模型的词典大小, 对于一个词 $w_i, x_i^{\text{Stat}}(d)$ 定义为:

$$x_i^{\text{Stat}}(d) = \begin{cases} f(w_i), & w_i \in d \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

其中, $f(w_i)$ 是词权重, 本文使用 $tf-idf$ 作为词权重。定义文档 d 的词向量特征表示:

$$\mathbf{x}^{\text{Vector}}(d) = (x_1^{\text{Vector}}(d), x_2^{\text{Vector}}(d), \dots, x_M^{\text{Vector}}(d)) \quad (9)$$

给定由 Neg-CBOW 模型训练得到的词向量矩阵:

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_M \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1N} \\ v_{21} & v_{22} & \cdots & v_{2N} \\ \vdots & \vdots & & \vdots \\ v_{M1} & v_{M2} & \cdots & v_{MN} \end{bmatrix} \quad (10)$$

得到 $\mathbf{x}^{\text{Vector}}(d)$ 的表达式:

$$\mathbf{x}^{\text{Vector}}(d) = \mathbf{x}^{\text{Stat}}(d) \mathbf{V} \quad (11)$$

将式 (11) 的文档 d 换成产品评论的标题和内容, 得到产品评论的词向量特征。

2.3 评估模型

本文使用支持向量回归模型 (SVR)^[14] 对评论的有用度进行评估。对于给定的训练数据 $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N) \subset X \times R$, 其中, X 表示输入向量的特征空间, SVR 模型尝试找到一个关于 y_i 的估计函数 $f(\mathbf{x})$, 使得对于所有输入 $\mathbf{x}_i, f(\mathbf{x})$ 与真实值 y_i 的偏差至多不超过 ε , 同时使得模型尽可能的扁平化。估计函数 f 的形式如下:

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \mathbf{w} \in X, b \in R \quad (12)$$

SVR 的优化目标函数为:

$$\min_{\mathbf{w}, \xi_i, \xi_i^*} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (13)$$

约束条件为:

$$\begin{cases} y_i - \mathbf{w}^T \mathbf{x} - b \geq \varepsilon + \xi_i \\ \mathbf{w}^T \mathbf{x} + b - y_i \geq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \leq 0 \end{cases} \quad (14)$$

3 实验与结果分析

产品评论有用度评估流程框架如图 2 所示, 主要步骤包括数据收集、数据预处理、特征抽取、模型训练、有用度评估与评论有用度排序。

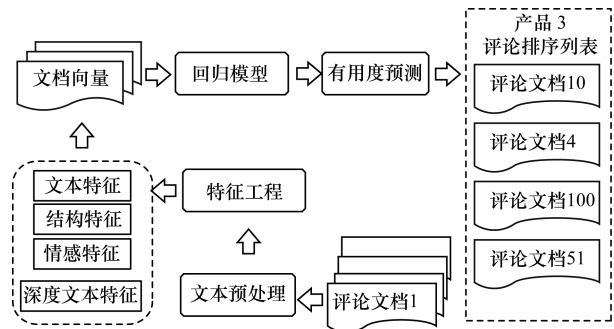


图2 基于词向量的产品评论有用度评估流程

3.1 数据集

本文研究的数据集为 Amazon.com 的产品评论

数据 MDSD (Multi-Domain Sentiment Dataset), 该数据集收集了 2006 全年发布于 Amazon. com 的大部分产品和评论并涵盖了 25 个产品类目, 比如 Books, Music, Video, Electronics 等。

参考前人工作的经验^[9], 本文采用下面 2 种方式对数据集进行过滤, 去除噪声与干扰:

1) 基于 Helpful Votes 的过滤

投票过少的评论, 不能客观地提供评论有用度标签, 本文设定有用度投票阈值 $\theta = 10$, 投票人数小于 θ 的评论不在考虑范围。

2) 基于 Bigram 的过滤

重复评论对于机器学习模型的训练起到负作用, 因此本文考察每一个评论对 (r_i, r_j) , 如果构成它们的 Bigram 有 80% 及以上的匹配, 则认为该评论对冗余, 仅保留其中一个。

经过以上的过滤, 本文最终得到 100 097 个产品以及 247 456 条评论。

3.2 特征抽取

3.2.1 文本深度特征

本文分别针对 3 个语料库进行词向量训练, 旨在挖掘领域相关 (domain dependency) 性对于词向量改善有用度评估模型的影响, 其中 Neg-CBOW 模型的负采样的参数设置为 3, 词向量大小设置为 300。

1) Google

由 Google News 语料库上训练的词向量模型, Google News 是比较正式的书面语文档集合, 由此训练的词向量模型反映了更加通用的语言特性。本文记该词向量模型表示的评论文档特征为 Google。

2) Amazon

Amazon Product Data 数据集收集了从 1996 年 5 月 - 2014 年 7 月间, 总计 1.42 亿的产品评论。本文抽取出其中每一条评论的评论内容和标题, 并过滤标点符号, 来构成领域相关的词向量训练语料。本文记由该词向量模型表示的评论文档特征为 Amazon。

3) MDSD

本文使用 Multi-Domain Sentiment Dataset 来训练领域相关性最高的词向量模型。在训练之前, 类似于 UGR 特征的提取, 本文对该语料库进行文本预处理, 包括去除标点符号、去除停用词、去除高频词、拼写纠错等。本文记由该词向量模型表示的评论文档特征为 MDSD。

3.2.2 传统 LEN, UGR, SENTI, STR 特征

本文使用 coreNLP 自然语言处理工具包, 依次对评论语料进行分词、词性标注、词形还原和命名实

体识别, 在此基础上进行相应的统计和计算, 生成 LEN, UGR, SENTI 3 组特征。此外, 本文使用 Jazzy 工具对语料进行了拼写纠错。STR 特征作为评论样本的元数据信息, 可以直接获得。

1) 结构特征

LEN——长度特征。本文对评论文档 (标题和内容) 进行句法解析, 抽取解析结果中的单词, 统计单词的长度以及句子的长度。

2) 朴素情感特征

SENTI——带有情感的评论更加具有可靠性 (credibility)。本文假设否定词 (don't, aren't, not, no 等)、疑问句、感叹句与产品评论的情感特征相关。本文统计否定词、疑问句、感叹句的数量作为评论的朴素情感特征。

3) 词汇特征

UGR——一元语言模型 (Unigram)。本文建立了一个去除停用词 (stopwords) 和低频词 ($df < 3$) 的词典, 评论标题和内容可以根据词典中的词 (term) 出现与否表示成一个稀疏的文档向量, 本文使用经典的 $tf-idf$ 权重为词赋权值并采用 $K = 0.5$ 的 2 倍归一化形式的 TF 权重:

$$K + (1 - K) \frac{f_{i,j}}{\max_k f_{i,j}} \quad (15)$$

和反比平滑文档频率形式的 IDF 权重:

$$\text{lb} \left(1 + \frac{N}{n_i} \right) \quad (16)$$

其中, $f_{i,j}$ 表示字词 i 在文档 j 中的频数; N 为语料中总文档数; n_i 为包含字词 i 的文档数目。

4) 元数据特征

STR——产品得分 (Star), 即评论作者对于产品的一个总体评价。

3.3 评价指标

本文使用多种评价指标来衡量基于词向量模型的有用度评估方法, 旨在证明词向量在各个方面能够改善产品评论有用度评估效果。本文使用 RMSE^[15] 评价有用度预测模型, 使用 Pearsons 系数、Spearman's 系数、Kendalls 系数以及 NDCG 度量来评估评论排序的效果。在计算 NDCG 时, 本文采用排序位置 i 来指定返回结果集中文档的分级相关得分, 其公式为:

$$\text{rel}_i = \frac{1}{\text{lb}i} \quad (17)$$

3.4 对比方法

本文对以下 5 种有用度评估方法进行了实验对比, 如表 1 所示。

表 1 实验对比方法

方法	特征组合
Kim	UGR + LEN + STR
Baseline	UGR + LEN + STR + SENTI
WV-Google	UGR + LEN + STR + SENTI + Google
WV-Amazon	UGR + LEN + STR + SENTI + Amazon
WV-MDSD	UGR + LEN + STR + SENTI + MDSD

在文献[5]中, Kim 使用 UGR、LEN 和 STR 3 种特征组合达到了令人满意的效果。Baseline 是本文提出的加入了朴素情感特征的基准方法。WV-Google 方法基于领域无关的词向量, WV-Amazon 和 WV-MDSD 方法基于领域相关的词向量。

3.5 实验结果

3.5.1 回归模型评价

如图 3 所示, 加入词向量特征, 对于评论有用度评估模型的效果均有较显著的提升。同时, 领域相关的词向量特征 (MDSD) 可以进一步降低 RMSE。

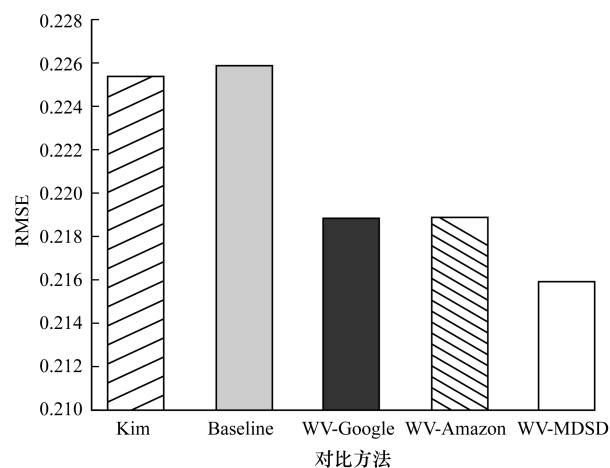


图 3 RMSE 结果对比 (越低越好)

3.5.2 排序评价

本文根据有用度评估模型对产品评论预测的有用度进行评论排序。表 2 展示了基于有用度评估模型的排序效果在 4 种评价指标上的实验结果。

表 2 在 4 个排序指标上的实验结果

Models	Pearsons				Spearman's			
	5	10	15	20	5	10	15	20
Kim	0.537 1	0.547 2	0.571 3	0.676 3	<u>0.481 7</u>	0.483 1	0.501 6	0.577 7
Baseline	0.535 4	0.547 7	0.572 2	0.671 8	0.479 8	0.486 2	0.506 0	0.587 3
WV-Google	0.537 4	0.554 9	0.573 4	0.725 7	0.474 7	0.492 5	0.505 1	0.647 7
WV-Amazon	0.538 9	0.557 7	0.577 5	0.724 0	0.477 0	<u>0.494 6</u>	0.520 3	0.660 6
WV-MDSD	<u>0.542 5</u>	<u>0.561 9</u>	<u>0.585 5</u>	<u>0.734 6</u>	0.477 5	0.494 3	<u>0.523 3</u>	<u>0.662 5</u>

Models	Kendalls				NDCG			
	5	10	15	20	@ 5	@ 10	@ 15	@ 20
Kim	<u>0.418 8</u>	0.404 7	0.411 9	0.479 9	0.801 4	0.811 3	0.819 8	0.837 2
Baseline	0.418 2	0.407 5	0.414 1	0.495 1	0.801 3	0.814 4	0.823 7	0.847 6
WV-Google	0.408 7	0.410 9	0.406 5	0.524 9	0.799 9	0.817 7	0.825 1	0.847 7
WV-Amazon	0.410 9	<u>0.414 1</u>	0.424 4	0.541 1	0.801 5	0.819 4	0.834 4	0.868 4
WV-MDSD	0.411 2	0.413 9	<u>0.427 7</u>	<u>0.556 3</u>	<u>0.803 3</u>	<u>0.822 2</u>	<u>0.835 1</u>	<u>0.869 7</u>

在表 2 中, 每列的最佳性能用下划线标示, 与 Baseline 或者 Kim 对比具有显著提升的用粗体标示。针对 Pearsons 系数、Spearman's 系数、Kendalls 系数这 3 种评价指标, 本文考察了评论列表长度大于等于 5, 10, 15, 20 这 4 种情况下各相关系数的情况。对于 NDCG 指标, 本文考察了 @ 5, @ 10, @ 15, @ 20 的情况。

实验结果表明, 在 4 种评价指标上, 相比于 Kim 提出的方法和本文提出的 Baseline 方法, 基于词向量的产品有用度评估模型均有比较明显的提升。此外, 实验结果证明了领域相关的词向量模型 Amazon 和 MDSD 与领域无关的词向量模型 Google 的效果相比, 具有进一步的提升, 这一点从图 4 可以更明显地看出。

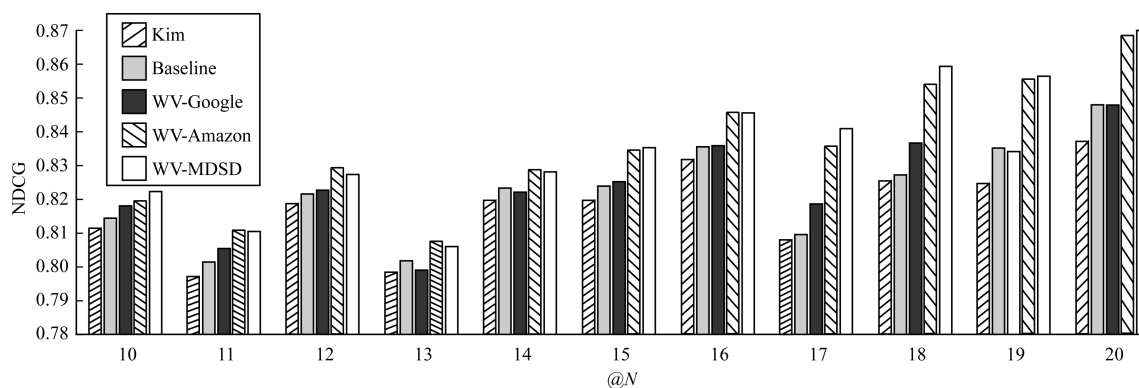


图 4 NDCG @ 10 ~ @ 20 结果对比

同时,MDSD方法与Amazon方法相比加入了领域相关性更高的词向量特征,实验结果表明,MDSD在诸多评价指标上均优于Amazon。

4 结束语

本文提出一种新的评估在线产品评论有用度的方法。新方法吸收了神经网络语言模型的思想,结合词向量模型,改善有用度评估的效果。本文使用真实数据集以及多种评价标准来评价基于词向量的有用度评估模型,证明了该模型的有效性。此外,本文探究了领域相关词向量模型与领域无关模型在产品评论有用度评估任务上的优劣性,通过多组实验对比,证明了领域相关的词向量可以更好地优化评估效果。下一步将侧重于训练更粗粒度的神经网络语言模型,生成句子向量或者文档向量。

参考文献

- [1] Jindal N, Liu Bing. Opinion Spam and Analysis [C]//Proceedings of the 2008 International Conference on Web Search and Data Mining. Los Angeles, USA: IEEE Press, 2008: 219-230.
- [2] Liu Jingjing, Cao Yunbo, Lin Chin-yew, et al. Low-quality Product Review Detection in Opinion Summarization [C]//Proceedings of EMNLP-CoNLL'07. Prague, Czech Republic: [s. n.], 2007: 334-342.
- [3] Collobert R, Weston J, Bottou L, et al. Natural Language Processing (Almost) from Scratch [J]. The Journal of Machine Learning Research, 2011, 12(1): 2493-2537.
- [4] 邱云飞, 王建坤, 邵良杉, 等. 基于用户行为的产品垃圾评论者检测研究 [J]. 计算机工程, 2012, 38(11): 254-257, 261.
- [5] Kim S M, Pantel P, Chklovski T, et al. Automatically Assessing Review Helpfulness [C]//Proceedings of 2006 Conference on Empirical Methods in Natural Language Processing. Sydney, Australia: [s. n.], 2006: 423-430.
- [6] 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.
- [7] O'Mahony M P, Smyth B. Using Readability Tests to Predict Helpful Product Reviews [C]//Proceedings of the 9th International Conference on Adaptivity, Personalization and Fusion of Heterogeneous Information. Paris, France: [s. n.], 2010: 164-167.
- [8] Liu Yang, Huang Xiangji, An Aijun, et al. Modeling and Predicting the Helpfulness of Online Reviews [C]//Proceedings of the 8th IEEE International Conference on ICDM. Pisa, Italy: [s. n.], 2008: 443-452.
- [9] Hong Yu, Lu Jun, Yao Jianmin, et al. What Reviews are Satisfactory: Novel Features for Automatic Helpfulness Voting [C]//Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval. Portland, USA: ACM Press, 2012: 495-504.
- [10] 孟佳娜, 段晓东, 杨亮. 基于特征变换的跨领域产品评论倾向性分析 [J]. 计算机工程, 2013, 39(10): 167-171.
- [11] Yang Yinfei, Yan Yaowei, Qiu Minghui, et al. Semantic Analysis and Helpfulness Prediction of Text for Online Product Reviews [C]//Proceedings of ACL'15. Beijing, China: [s. n.], 2015: 38-44.
- [12] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and Their Compositionality [C]//Proceedings of ANIPS'. Lake Tahoe, USA: [s. n.], 2013: 3111-3119.
- [13] Mikolov T, Chen Kai, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space [C]//Proceedings of ICLR'13. Scottsdale, USA: [s. n.], 2013: 254-262.
- [14] Smola A J, Schölkopf B. A Tutorial on Support Vector Regression [J]. Statistics and Computing, 2004, 14(3): 199-222.
- [15] Rob H, Koehler J, Anne B. Another Look at Measures of Forecast Accuracy [J]. International Journal of Forecasting, 2006, 22(4): 679-688.

编辑 索书志

(上接第187页)

- [9] Nelder J A, Baker R J. Generalized Linear Models [M]. [S. l.]: Wiley Online Library, 1972.
- [10] Dempster A P, Laird N M, Rubin D B, et al. Maximum Likelihood from Incomplete Data via the EM Algorithm [J]. Journal of the Royal Statistical Society, 1977, 39(1): 1-38.
- [11] Ripley B D. Pattern Recognition and Neural Networks [M]. [S. l.]: Cambridge University Press, 2007.
- [12] Rätsch G, Onoda T, Müller K R. Soft Margins for Adaboost [J]. Machine Learning, 2001, 42(3): 287-320.
- [13] Blake C, Merz C J. UCI Repository of Machine Learning Databases [2011-11-23]. <http://www.ics.uci.edu/mlearn/Machine-Learning.html>.
- [14] Alon U, Barkai N, Notterman D A, et al. Broad Patterns of Gene Expression Revealed by Clustering Analysis of Tumor and Normal Colon Tissues Probed by Oligonucleotide Arrays [C]//Proceedings of National Academy of Sciences, 1999, 96(12): 6745-6750.
- [15] Cawley G C, Talbot N L C, Girolami M. Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation [C]//Proceedings of the 20th Annual Conference on Neural Information Processing Systems. Vancouver, Canada: [s. n.], 2007: 209-216.

编辑 索书志