



数据分析与知识发现  
*Data Analysis and Knowledge Discovery*  
ISSN 2096-3467, CN 10-1478/G2

## 《数据分析与知识发现》网络首发论文

题目：基于深度学习的众测报告有用性预测研究  
作者：蔡婧璇，吴江，王诚坤  
网络首发日期：2020-09-29  
引用格式：蔡婧璇，吴江，王诚坤. 基于深度学习的众测报告有用性预测研究. 数据分析与知识发现.  
<https://kns.cnki.net/kcms/detail/10.1478.G2.20200928.1917.004.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于深度学习的众测报告有用性预测研究\*

蔡婧璇<sup>1</sup>，吴江<sup>1, 2</sup>，王诚坤<sup>1</sup>

<sup>1</sup>(武汉大学信息管理学院 武汉 430072)

<sup>2</sup>(武汉大学电子商务研究与发展中心 武汉 430072)

## 摘要：

**[目的]** 本文以众测报告为研究对象，探索众测报告作者属性、产品属性、文本、图片对预测众测报告有用性的作用。

**[方法]** 本文使用深度学习提取众测报告的文本特征和图片特征，使用全连接神经网络构建众测报告有用性预测模型，使用 80%随机样本对不同输入组合下的模型进行训练学习，再用剩余样本作为测试集评估模型的预测效果。

**[结果]** 单独加入文本特征后，模型的预测效果提升 5.430%；单独加入图片特征后，模型的预测效果提升 6.659%；同时加入文本特征和图片特征后，模型的预测效果提升 8.895%；

**[局限]** 深度学习提取的文本特征和图片特征可理解性和可解释性较差，因此，即使最终模型的预测效果比较准确，仍难以得知模型中每一层神经网络所代表的具体特征并总结归纳出模型做出最终的决策结果所依赖的预测规则。

**[结论]** 众测报告中文字描述的特征和图片的特征都能有效预测众测报告对消费者的有用性，且这两者对于预测众测报告对消费者的有用性上具有相互验证和相互替代的作用。

**关键词：** 产品众测；信号理论；深度学习；特征提取；预测分析

**分类号：** G203

**DOI:** 10.11925/infotech.2096-3467.2020.0059

---

(通讯作者：吴江，ORCID：0000-0002-3342-9757，E-mail:jiangw@whu.edu.cn。)

\*本文系国家自然科学基金项目“信息不对称驱动的共享经济去中心化机制与风险的复杂性研究”（项目编号：71874131）的研究成果之一。

# Research on the Prediction for the Usefulness of Crowd Testing Reports Based on Deep Learning

Cai Jingxuan<sup>1</sup>, Wu Jiang<sup>1,2</sup>, Wang Chengkun<sup>1</sup>

<sup>1</sup>(School of Information Management, Wuhan University, Wuhan 430072, China)

<sup>2</sup> (Center for E-commerce Research and Development, Wuhan University, Wuhan 430072, China)

## Abstract:

**[Objective]** This paper takes the crowd testing reports as the research object to explore the prediction for the usefulness of crowd testing reports with author attributes, text features, and image features.

**[Methods]** In this paper, deep learning was adopted to extract text features and image features of reports. The full-connected neural network was used to construct a prediction model for the usefulness of crowd testing reports. 80% of samples were used to train models with different input combinations. The remaining samples serve as test sets to evaluate the prediction effect of the model.

**[Results]** When text features were added, the prediction effect of the model increased by 5.430%. When image features are added, the prediction effect of the model increased by 6.659%. After adding both text features and picture features, the prediction effect of the model increased by 8.895%.

**[Limitations]** Text features and image features extracted by in-depth learning are poorly understandable and interpretable. Therefore, even if the prediction effect of the final model is relatively accurate, we still have no cue for the specific features represented by each layer of neural network in the model and summarize the prediction rules that the model relies on to make the final decision.

**[Conclusions]** The features of text and image can effectively predict the usefulness of crowd testing reports to consumers. And these two can verify and replace each other in predicting the usefulness of crowd testing reports.

**Keywords:** Crowd Testing; Signal Theory; Deep Learning; Feature Extraction; Predictive Analysis

# 1 引言

随着电子商务的蓬勃发展，商品免费众测已经成为商家用来提高产品知名度和涉入度的常见营销方法<sup>[1, 2]</sup>。商品众测是指商家免费向选定的客户发放试用品，用户在试用后以报告的形式提交反馈以供广大消费者参考的一种营销手段。对商家而言，众测可以提升顾客对产品的接受度，从而提高其销量。同时，也能提高产品的网络口碑。更重要的是，这种营销形式是顾客喜闻乐见的。在强调客户体验的市场中，消费者参与免费的产品众测，能大大减少消费者在线上购物中面临的风险。对于获得产品试用机会的申请者来说，他们可以免费获得商品或者服务，提前感知到产品的使用体验<sup>[3]</sup>。尽管被接受的试用申请数量有限，并非所有消费者都能收到产品，但是即使没有获得试用机会，他们也能从他人的众测报告中获得详实的使用体验，对产品形成更全面深入的了解，从而提高购买的可能性<sup>[4]</sup>。

产品众测与传统的产品试用有许多相似之处，但也有很大的不同。首先，面向群体的广度不同。传统的产品试用仅仅针对参与产品试用的人群，使试用者对产品产生好感，从而产生消费的需求<sup>[5]</sup>。产品众测可以通过报告的形式，将真实的产品体验传递给广大消费者。众测报告作为一种网络口碑，其后续带来的影响及宣传效果（该产品的品牌、商家及其他相关产品）能波及到整个平台的消费者群体<sup>[6]</sup>，其数量和规模都是产品试用不可比拟的。其次，两者的试用流程不同。产品试用是由商家主动将试用品随机发放或随商品附赠给商家认为有潜在消费需求的顾客。而产品众测是在商家发布众测信息后，由众测平台用户提出申请，商家或平台进行筛选后公布最终入选名单，商家将产品邮寄给被选中的用户，用户试用体验商品后提交测试报告，最终将测试报告发布与平台上供广大消费者参考<sup>[6]</sup>。由于申请流程需要填写个人信息及申请理由等资料，并通过商家和平台的筛选得出最终试用名单，这一过程排除了许多投机的申请者，极大地保证了试用申请者多是对产品有需求、感兴趣的消费者<sup>[7]</sup>。因此，无论在辐射广度和宣传效果上，众测的营销效果较传统的产品试用都有较大提升。

传统的计算机程序通过人们硬编码而实现的，对执行任务使用特定的指令，但这种程序灵活度不够，难以处理现实世界中的复杂问题。深度学习利用标注好的训练数据为计算机程序提供了大量的经验，从中学习出基本原则以便识别数据中蕴含的模式、关联和见解<sup>[8]</sup>。因此，深度学习能够从复杂数据中自动提取层次特征，被广泛应用于模式识别、分类和预测问题<sup>[9, 10]</sup>，已经成为目前各研究领域研究热点。深度学习利用多层神经网络结构，旨在从大数据中学习出现实世界中各类事物能直接用于计算机计算的表示形式，对于图片、文

本、声音等无法直接抽象出有效特征的数据，依赖于使用神经网络进行多个层次的转换来提取特征<sup>[11, 12]</sup>。首先提取较低层次的简单特征，再基于这些简单特征来表示更复杂、更抽象的特征<sup>[13]</sup>。通过逐层特征提取，将数据在原始空间的特征转换到一个新的特征空间来表示初始数据，随着网络深度的增加，充分的特征转换将简单的特征合并到复杂特征中，从原始数据中提取出最合适的层次表示，使得分类或预测问题更加容易实现<sup>[14]</sup>。

总结而言，目前对产品众测的研究较少，且主要集中在产品众测对产品销量增长的促进作用。然而，鲜有研究以众测报告为研究对象，探索众测报告对消费者有用性的研究。为此，本文以平台用户对众测报告的点赞数作为众测报告有用性高低的指标，使用深度学习的方法，构建众测报告的作者属性、产品属性、文本、图片对众测报告有用性的预测模型，研究及评估影响众测报告有用性的重要特征。

## 2 理论基础与研究假设

### 2.1 产品众测研究

对于产品众测的研究，学者更多考虑的是研究的实践意义，因此，研究多聚焦于试用营销的效果或新产品的市场开发方面。全新的商品或服务在进入新的市场时，消费者将感知到两种风险：产品功能的不确定性和产品质量带来的使用体验。于消费者而言，产品试用可以节约时间成本和金钱成本，减少不确定性和风险，通过产品试用获得的产品信息判断商品价值<sup>[15]</sup>。同时，消费者在进行产品试用时涉入度更高，试用报告中的直接体验比间接的广告信息更能影响消费者对产品的认知<sup>[16]</sup>。试用报告作为深度的网络口碑，能促进消费者的购买决策，并且受到消费者信任的调节作用。产品试用报告和常规评论结合，可以促进体验性产品的销售，但试用报告描述过于正面时，会使消费者怀疑其客观性。

虽然消费者对于产品众测这一营销方式的接受度较高，但仍不可避免会产生一些问题。试用营销在带来积极的网络效应的同时，也会产生蚕食效应。这体现在产品试用可以增加产品的用户基础，提升产品网络口碑，但营销所提供的试用产品也会蚕食部分正式产品的市场，即消费者的需求已经从试用产品中得到满足而不再需要购买正式产品。

已有的产品试用营销相关文献，主要论证了试用的有效性<sup>[17, 18]</sup>。试用营销可以有效降低消费者的感知风险，特别是在新产品和新市场的开发中发挥作用。免费试用对于消费者对产品的态度和购买意愿都有显著的影响<sup>[19]</sup>。然而，



目前的研究多是关注这种营销手段对促进消费的有效性。以消费者为对象，研究众测报告本身对消费者有用性影响的研究尚少。因此，本文为了弥补这一研究缺口，从众测报告入手，使用深度学习的方法提取报告中文本特征和图片特征，结合作者属性和试用产品属性等结构化特征，以消费者的点赞行为为目标，预测众测报告的有用性。

## 2.2 信号理论

信号理论最初来源于信息经济学的领域，被广泛应用于管理学、营销学、医患关系等领域中分析信息不对称性的问题<sup>[20-22]</sup>。信号理论认为，处于信息不对称场景中的用户在评判产品或服务时主要采取信号甄别和信号传递两种途径。信息甄别是指信号的接收方通过各种信号评估甄别信号发送方的真实信息；信号传递是指信号发送方通过可观察的行为传递商品价值或质量的确切信息<sup>[20]</sup>。

由于网络环境中的信息不对称，信号对于衡量潜在的、不可观察的产品质量、可信度和真实性起着重要作用<sup>[23]</sup>。传统实体店交易中，产品在消费者挑选过程中可以被直接接触和观察，消费者对产品质量的把控也更加清晰直观。而在产品众测的背景下，消费者无法直接对商品或服务本身进行体验，更多地只能依赖于试用者对商品体验的描述。

## 2.3 研究假设

基于信号理论，消费者无法感知产品的全部质量信息和众测报告试用者的可靠度，增加了其感知产品质量的不确定性。尤其对于服装或体验类产品等非电子属性的产品，由于众测报告中关于产品物理线索的缺失，消费者只能依赖于报告中展示的信号作为产品质量的诊断线索<sup>[24]</sup>。信号则是用来解决这些问题的，它是在线交易中消费者决策过程的主要信息来源，具有重要的预测价值。在网络交易中，通过特定网站功能的展现，包含着产品信息的信号在众测报告中以文字描述和图片的形式被传递给买方。因此，本文认为：

H1: 众测报告中文字描述的特征能有效预测众测报告对消费者的有用性。

H2: 众测报告中图片的特征能有效预测众测报告对消费者的有用性。

信号理论的一个核心观点在于不同类型的信号所蕴含的信息量不同，在不同使用场景下发挥的作用也不同<sup>[25]</sup>。它解释了信号与产品质量之间的关系，说

明了为什么一些信号是可靠的而另一些不可靠<sup>[26]</sup>。因此，本文关注了产品众测报告中作者属性、产品属性、报告文本属性、报告图片属性这四类信号对房源销量的影响程度，并对比了基于视觉的图片信号和基于文本的文字描述信号对众测报告有用性的边际信息价值。因此，本文认为：

H3: 众测报告中文字描述的特征和图片特征在影响众测报告对消费者的有用性上具有相互替代的作用。

### 3 研究方法

#### 3.1 研究对象与数据获取

本文选取全球领先的导购平台 Dealmoon 作为研究对象。Dealmoon 于 2009 年在美国推出，是为在世界各地的华人提供购物建议的社交购物社区。2016 年 9 月，Dealmoon 推出了产品众测服务。最初，参与众测的人员主要来自于 Dealmoon 团队工作人员，所有众测报告都由官方账号发布。2017 年 2 月，为了增加用户参与度，提高用户参与体验，产品众测完全向公众开放，网站的所有用户都可以申请参与众测并发布报告。随后，报告变得更多样化，具有试用人员明显的个人风格。

本文使用 Python 爬虫获取 Dealmoon 上全部众测报告，包括美妆护肤、美食、生活、服饰、电子产品、母婴、旅行和汽车 8 个种类。由于旅行和汽车的众测与其他种类实行方式不同，且所占比例较小，本文中剔除这两个分类下的众测报告。在进行数据清洗后，获得 1550 份众测报告。

#### 3.2 数据处理与特征提取

在本文的研究数据集分为结构化数据和非结构化数据。结构化数据包括众测报告点赞数、作者属性和产品属性，非结构化数据包括众测报告文本和图片。表 1 展示了包括众测报告点赞数、作者属性和产品属性在内的结构化基础属性定义及描述性统计。用户在阅读众测报告后可以对认为有用的报告进行点赞，因此本文使用众测报告获得的点赞数 *Like* 作为其有效性的衡量指标。作者属性是指在众测报告页面可见的与报告作者相关的信息，主要包括作者等级 *Au\_level*、作者在平台上发布所有文章数 *Au\_article*、作者参与众测次数 *Au\_test*、作者关注数 *Au\_followed*、作者粉丝数 *Au\_follower*。产品属性包括产品类型 *Prod\_type*、产品价格 *Prod\_price*、同时参与该商品众测人数 *Prod\_num*。

表 1 基础属性定义及描述性统计

Table1 Variables Definitions and Descriptive Statistics							
维度	变量	说明	Obs	Mean	S.D.	Max	Min
因变量	<i>Like</i>	众测报告获得的点赞数	1550	27.281	33.558	385	0
作者属性	<i>Au_level</i>	作者在众测平台上的等级	1550	9.545	3.722	26	1
	<i>Au_article</i>	作者在平台上发布的文章总数	1550	24.568	31.184	246	1
	<i>Au_test</i>	作者参与众测次数	1550	5.097	4.816	114	1
	<i>Au_followed</i>	作者关注数	1550	911.739	1486.366	10620	0
	<i>Au_follower</i>	作者粉丝数	1550	57.021	66.684	1225	0
产品属性	<i>Prod_type</i>	试用产品的类型（美妆护肤、美食、生活、服饰、电子产品、母婴）	1550	--	--	--	--
	<i>Prod_price</i>	试用产品的价格	1550	177.319	215.297	1700	11.99
	<i>Prod_num</i>	同时参与该商品众测人数	1550	7.982	7.348	32	1

众测报告的作者属性和测试产品属性是结构化数据。与图片和文本的高维特征表示不同，有关试用者属性的结构化数据基本特征维度简单，信息组合和丰富度较差。因此本文还提取了众测报告的文本特征和图片特征。

本文首先使用 Jieba 分词工具对众测报告的文本进行分词处理。Jieba 是目前运用最广泛的中文词法分析器，其最新的分词系统分词精度高达 98.45%。由于该平台属于提供购物建议的社交购物社区，文本中包含大量有关互联网、产品及商家名称、产品推荐等的专业词汇和新兴词汇，为使分词得到更好的效果，本文建立了基于评论语境的用户词典和停用词典，具体如表 2 所示。用户词典包括产品及商家名称、体验描述类用词和网络用语及缩写，作用是为了对购物建议类的专属词进行识别，达到更准确的分词效果。停用词典来自于哈工大停用词表，该词表是中文常用的停用词表，作用是过滤掉分词结果中无意义的停止词，如标点符号、虚词、表情及其他无意义的词语。最后，为了进一步保证分词的完整性，本文对分词后的结果进行词频统计，对词频大于 100 的词进行人工核对。无意义词纳入停用词典，将误分的词纳入用户词典中。



表 2 用户词典及停用词表

Table2 User Dictionary and Stop Words

类别	内容	来源
用户词典	1. 产品及商家名称：美妆蛋、小白鞋、良品铺子等	1. 网站中对众测商品的描述
	2. 体验描述：沙漠皮、开箱、神器等	2. 高频未分词
	3. 网络用语及缩写：鸡冻、踩雷、轻奢等	
停用词典	1. 标点符号	1. 哈工大停用词表
	2. 虚词	2. 高频无意义词
	3. 表情	
	4. 其他无意义词语：哈哈等	

本文对文本内容的分词结果进行 TF-IDF 计算，并进一步计算各主题分布下的困惑度。得到的困惑度计算结果如图 1 所示。通过对比主题数的增长和困惑度的降低，确定了 LDA 模型参数  $\alpha=0.05$ ，主题参数  $K=20$ ，为该模型的最佳主题数。最终得到每篇报告在 20 个主题的分布概率，作为文本特征加入最终模型进行预测。

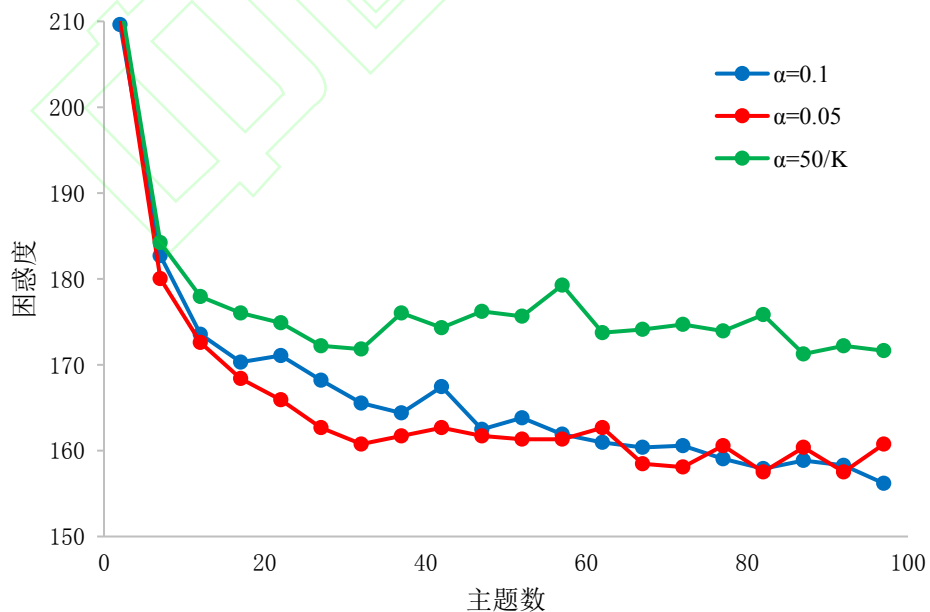


图 1 众测报告文本主题困惑度

Fig. 1. Topic Perplexity of Crowd Testing Report

本文使用 VGG19 图片分类模型来对报告中的图片进行特征提取，共包括 19 个卷积层、5 个最大池化层和 4 个全连接层。VGG19 网络通过多个 3×3 的卷积核来替代早期的大尺寸卷积核，大大减少了参数的数量，提高了运算效率。VGG19 模型的输入为尺寸为 224×224×3 的图像，采用 3×3 大小的卷积核和大小为 2×2、步长为 2 的最大池化，通过层次学习对图像进行特征提取。原始的 VGG19 模型通过 19 个卷积层来进行图像的特征提取，使用 4 个全连接层对图像进行分类，但其分类目标是对图片进行物体识别的分类，而本文中的分类目标是基于图像特征以及其他特征，对众测报告获得点赞数的高低进行分类，因此本文没有采用 VGGNet 模型自带的全连接层，而只利用其 19 个卷积层和 5 个最大池化层来进行图像特征的提取。如图 2 所示，本文中采用众测报告中呈现的第一张带有标题的主图作为提取图片特征的原始数据，使用预训练的 VGG19 对图片进行特征提取。本文通过网页爬虫采集到的房源图片原始大小为 1000×562×3，将房源图片转换为 224×224×3 的尺寸输入到 VGG19 模型中，使用 19 个卷积层进行特征提取，最后一层卷积网络输出的 7×7×512 尺寸的特征图即为所提取出的图片特征，这其中可能包含了图片的纹理、色彩、明暗、物体等特征。

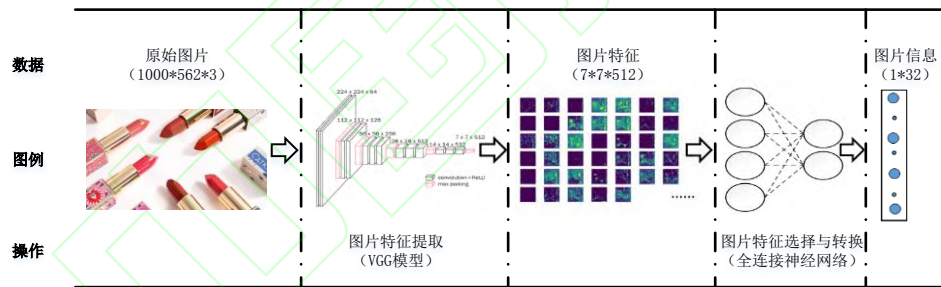


图 2 图片特征提取流程

Fig. 2. Image Feature Extraction Process

3.3 模型验证

为了验证报告文本和图片对预测众测报告有用性的影响程度，本文以作者属性和产品属性作为基本特征，将图片特征和文本特征作为附加特征，通过组合的方式形成 4 种有用性预测模型：模型 1：只包含作者属性和产品属性的基本特征预测模型；模型 2：基本特征+文本特征的众测报告有用性预测模型；模

型 3：基本特征+图片特征的众测报告有用性预测模型；模型 4：基本特征+图片特征+文本特征的众测报告有用性预测模型。并以基本特征预测模型效果为基准来评价加入其它特征后的预测模型效果提升程度，对比报告文本和图片对模型准确性的影响。

本文使用深度学习的方法提取众测报告的各种特征，预测众测报告对浏览者的有用性。为了提升模型的预测效果，本文以高点赞和低点赞作为感知有用性的指标，将每篇报告所获得的点赞数转化为高点赞/低点赞的二分类标签，作为模型的输出。选取了 1550 篇众测报告中获得点赞数的前 40%（620 篇）和后 40%（620 篇），转化为高点赞/低点赞的二分类标签，作为模型的输出。分别对报告文本和报告图片进行特征提取，与其他结构化特征（作者属性和产品属性）利用全连接神经网络进行特征融合。以包含全部基本特征、文本特征、图片特征的模型 4 为例，本文通过对基本特征、文本特征、图片特征分别转换为 32 个维度的特征，在全连接神经网络中进行叠加，形成 96 个维度的特征，再进行后续训练。不同特征的融合过程如图 3 所示。此后，构建分类模型，对众测报告的有用性进行预测。模型图如图 4 所示。

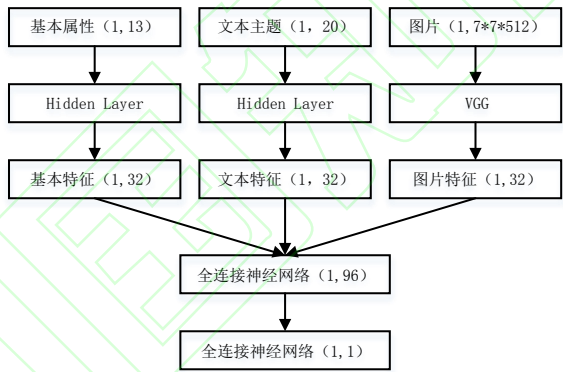


图 3 特征融合流程图（以模型 4 为例）

Fig. 3. Feature Integration (Model 4)

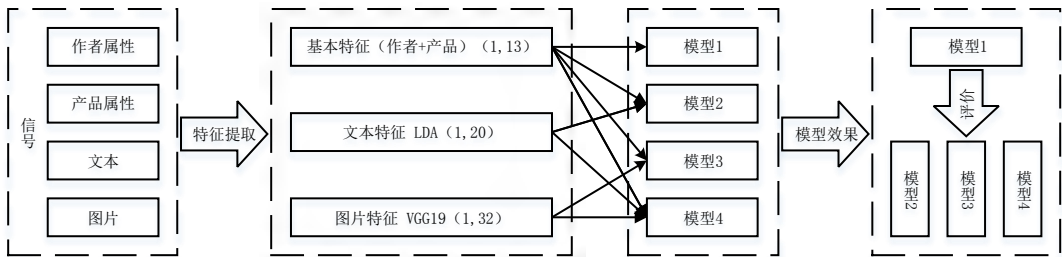


图 4 模型构建流程图

Fig. 4. Model Construction

实验过程中，随机抽取 1000 条记录作为训练集，用于分类模型的训练，将其余 240 条记录作为测试集，用于验证模型的准确性。同时，为了防止出现过拟合现象，本文采取了交叉验证方法，对原始数据集进行多次随机打乱，重新分割出训练集和测试集，用于训练和测试模型的准确性。最终，本文对 4 种输入组合的模型下，均通过 10 次交叉验证对模型准确性进行了训练和测试，最终得到每种输入组合下所训练模型在 10 次交叉验证后的平均准确性和性能。

## 4 研究结果

本文基于众测报告的作者属性、产品属性、报告文本、报告图片等特征，采用深度学习模型，来预测该众测报告对消费者有用性的高低。通过对比不同类型特征的输入下模型预测效果的提升程度，来获取报告文本和图片特征对预测效果的影响，作为文本和图片对模型效果带来的边际效益。本文中深度学习模型的实现主要分为两大步骤：首先，使用预训练的 VGG16 模型对报告图片进行图像特征提取，利用 LDA 模型提取文本主题的特征表示，并结合作者属性和试用产品属性的特征表示共同构成用于预测的序列输入特征；然后针对不同的模型输入组合，试用全连接神经网络进行序列建模，并对模型进行训练和测试。

基于训练好的 4 个模型，本文使用事先预留的测试集用于对模型效果的测试，并得到模型对测试集的预测结果。在模型预测效果的评估上，本文使用精确率（Precision）、召回率（Recall）、F1 值和准确率（Accuracy）指标进行评估。这 4 种指标是针对二分类问题的典型的算法效果评价指标，被广泛应用于深度学习、机器学习和信息检索等领域。不同类型特征组合下的 4 种模型对测试集的预测效果如表 3 所示。

本文中各个模型的精确率、召回率及 F1 值均较高，因此本文构建的众测报告有用性预测模型是可靠的。其中，模型 1 仅包含基础的结构化特征即作者属性和产品属性，模型 2、模型 3 分别加入了众测报告的文本特征和图片特征，模型 4 在模型 1 的基础上将报告文本和图片结合起来，加入了文本和图片两种类型的特征。因此将模型 1 作为基础模型，通过对比其他模型相对模型 1 在预测效果上的提升程度，可以得知不同类型特征对模型的提升影响。

表 3 模型预测效果

Table3 Model Performance

模型代码	特征组合	精确率	召回率	F1 值	准确率
模型 1	作者属性+产品属性	0.7676	0.8783	0.7917	0.7821
模型 2	作者属性+产品属性+报告文本	0.8624	0.8403	0.8099	0.8245
模型 3	作者属性+产品属性+报告图片	0.86996	0.8571	0.8359	0.8342
模型 4	作者属性+产品属性+报告文本+报告图片	0.8854	0.8870	0.8658	0.8517

为了进一步提升模型的性能，本文还使用近年来常用的 ResNet50、InceptionV3、NASNetMobile 图片预训练模型进行特征提取。并用不同模型提取出的特征再次运算模型 3 和模型 4，结果如表 4 所示。模型 3 中使用 VGG19 进行图片特征提取的结果准确率为 0.8342，好于 ResNet50（0.8167）、InceptionsV3（0.5123）、NASNetMobile（0.7688）。模型 4 中使用 VGG19 进行图片特征提取的结果准确率为 0.8517，好于 ResNet50（0.7938）、InceptionsV3（0.8313）、NASNetMobile（0.8105）。对于精准率、召回率、F1 值等其他指标，除使用 InceptionV3 图片预训练模型的模型 3 多次出现过拟合现象外，使用 VGG19 图片预训练模型的预测效果也优于使用 ResNet50、InceptionsV3、NASNetMobile 的预测效果。由此可见，VGG19 图片预训练模型在处理本文中预测任务的整体指标较其他模型好。因此，本文将采取 VGG19 的结果进行下一步分析。



表 4 使用不同图片预训练模型的预测效果

Table4 Model Performance with Different Picture Pre-training Model

模型代码	图片预训练模型	精确率	召回率	F1 值	准确率
模型 3	VGG19	0.86996	0.8571	0.8359	0.8342
	ResNet50	0.8430	0.8031	0.8226	0.8167
	InceptionV3	0.5041	0.9419	0.6701	0.5123
	NASNetMobile	0.7952	0.7923	0.7790	0.7688
模型 4	VGG19	0.8854	0.8870	0.8658	0.8517
	ResNet50	0.7860	0.8381	0.8036	0.7938
	InceptionV3	0.8598	0.8067	0.8316	0.8313
	NASNetMobile	0.8850	0.7455	0.8011	0.8105

4.1 文本特征对预测众测报告有用性的影响

模型 1 的精确率、召回率、F1 值以及准确率分别为 0.7676、0.8783、0.7917、0.7821。模型 2 的精确率、召回率、F1 值以及准确率分别为 0.8624、0.8403、0.8099、0.8245。以模型 1 为基础模型，模型 2 的表现均优于模型 1，在加入报告文本后，模型的准确率从 0.7821 提高至 0.8245，说明报告文本可以提高模型对众测报告有用性的高低的预测的准确率。通过众测报告中的文字描述，消费者可以获得有关产品的更直观和详细的信息，作为最终决策的辅助，增强对产品或服务质量的诊断能力，降低用户的感知风险和不确定性。因此，报告文本的特征可以作为反映众测报告质量的一个信号，是反映其有用性的一个重要因素。

4.2 图片特征对预测众测报告有用性的影响

模型 3 的精确率、召回率、F1 值以及准确率分别为 0.86996、0.8571、0.8359、0.8342，均高于模型 1，加入图片特征后模型的准确率为 0.8517，相比基础模型也有所提升，表明报告中的图片特征也有助于提高模型的预测效果。

图片提供的视觉吸引力对吸引用户注意力和影响用户决策发挥了重要作用，因此高质量的产品试用图片对众测报告的有用性也会产生正向促进作用<sup>[27, 28]</sup>。

同时，模型 3 的预测效果在各个指标上的表现均优于模型 2，表明相比报告中的文字描述，加入图片特征对基础模型的准确率的提升效果更好。一方面是因为，消费者对不同类型的信息通常采用不同的信息加工策略，即经验系统和分析系统<sup>[29]</sup>。图片为用户提供了直观的视觉线索，用户通过观察图片获取的信息和产生的情感是感性的、直觉的、和自动的感知，因此通过经验系统即可进行快速的评估与决策；而文字描述需要用户利用分析系统进行理性的、分析的信息加工。而在感官型和体验型产品场景下，用户会优先使用与直觉相关的经验系统进行信息加工，因此在感知众测报告有用性的场景下图片信号会比文本信号更加重要。另一方面，视觉信号所引起的图像信息处理过程通常是产生对图片中物体的想象，包括了如视觉、味觉、触觉、嗅觉等多个感官维度<sup>[30]</sup>，而用户对文本信号的处理往往缺乏对信息的感官体验，因此用户对文本信息的感知不如图像感知更具体，从而导致图像信息对用户的影响高于文本信息<sup>[31]</sup>。

### 4.3 文本特征与图片特征对预测众测报告有用性的相互作用

相比于基础模型，包含了全部类型特征的模型 4 在精确率（0.8854）、召回率（0.8870）、F1 值（0.8658）以及准确率（0.8517）上均较高，同时加入文本和图片特征后模型的准确率相比基础模型提高了 8.895%。模型 4 在各个指标上的表现均优于模型 2 和模型 3，表明将众测报告的文本特征和图片特征结合起来用于预测众测报告对消费者的有用性相比分别将这两类特征单独用于模型预测的效果更好。说明文字描述和图片在向消费者传递信息方面可以相互补充、验证和增强，说明在体验产品场景下，基于图片的信号和基于文本的信号对消费者的态度具有叠加效应。

但同时，模型 4 的准确率相对模型 2 仅提升了 3.287%，相对模型 3 仅提升了 2.097%，准确率的提升程度小于模型 2 和模型 3 相对模型 1 的提升程度。因此，虽然将文本和图片结合起来共同用于众测报告有用性高低的预测效果更好，但对模型准确率的提升程度较小，这说明文本和图片之间可能存在一定的信息冗余，在蕴含的信息量上具有一定的可替代性。通过阅读报告中的文字描述和浏览图片，用户就可以对产品或者服务的质量进行较为准确的感知和预估。但不可否认的是，文本和图片之间确实有着相互补充、加强和验证的作用<sup>[32]</sup>。在文字描述的基础上，图片对产品信息进行了补充和验证，进一步提高了该众测报告的可信度和吸引力。

## 5 研究结论与讨论

本文以 Dealmoon 产品众测平台为研究对象，收集了平台上 1550 份产品众测报告的相关信息，包括点赞数、作者属性、试用产品属性、报告文本和报告图片，使用深度学习的方法，提取众测报告中文本特征和图片特征对众测报告有用性进行预测。结果表明，众测报告中文字描述的特征和图片的特征都能有效预测众测报告对消费者的有用性，且这两者对于预测众测报告对消费者的有用性上具有相互验证和相互替代的作用。

从理论上，对比了基于视觉的信号（图片）和基于文本的信号（文字描述）对信息有用性的影响程度，并对比了这两种信号在单独使用和组合使用下对有用性的影响。研究结果表明加入图片特征对模型准确率的提升效果高于加入文本特征后的提升效果，表明基于视觉的图片信号相比于基于文本的信号在影响用户决策上有更重要的作用<sup>[30]</sup>。并且，同时加入文本特征和图片特征后模型准确率均高于分别加入文本和图片的模型准确率，表明文字信号和图片信号在传递信息方面可以相互补充、验证和增强，且基于图片的信号和基于文本的信号之间对用户的态度具有叠加效应。但同时加入文本和图片特征后模型准确率相比于单独加入文本或图片的模型准确率的提升程度相对较小，没有产生  $1+1>2$  的效果，说明文本和图片在蕴含的信息量上存在一定的替代性。这可以理解为，众测报告中的文字描述和图片在体现产品细节的信息量上既有独立的、相互补充的部分，也存在一定的信息冗余，起到了相互验证和增强的作用<sup>[32]</sup>。本文的研究结论为产品图像与文字描述之间的关系做出了阐释，未来的研究可进一步挖掘在不同背景下这两类信号对用户的影响，以及研究图像、文本以及其他信号之间的相互作用机制，以便更全面地理解不同信号之间如何相互影响。

同时，本文也对众测平台的管理和报告作者具有重要的实践指导意义。本文明确了在众测报告中展示的图片对报告的有用性有着显著的提升作用，对消费者来说，图片是判断商品质量的一个重要线索<sup>[33]</sup>。因此，众测报告的作者要对报告中图片的拍摄美感和图片质量引起足够的重视。对于众测平台而言，为了给消费者更好的服务体验，可以从提高众测报告文字描述和图片质量两个方面来提升众测报告对消费者的有用性，例如为报告作者提供有关众测报告撰写的指导。

然而，本文还存在一些局限。虽然深度学习能够通过多层神经网络自动提取图片、文本中的特征，但其可理解性和可解释性较差。因此，即使最终模型的预测效果比较准确，我们仍无法得知模型中每一层神经网络所代表的具体特征并总结归纳出模型做出最终的决策结果所依赖的预测规则。即，虽然最终模型的预测结果说明加入文本和图片的特征有助于提高模型的预测准确率，但我

们无法识别出文本和图片中的哪些元素、特征或由图片反映出的产品的哪些方面对预测结果起到了重要的作用。而识别图片中的重要元素和特征对在众测平台的管理和报告作者来说具有很重要的实践指导意义，因此这是未来进一步的研究方向之一。

### 参考文献:

- [1] HAN Y, ZHANG Z. Impact of free sampling on product diffusion based on Bass model [J]. Electronic Commerce Research, 2018, 18(2): 1-17.
- [2] KOCH O F, BENLIAN A. The effect of free sampling strategies on freemium conversion rates [J]. Electronic Markets, 2017, 27(1): 1-10.
- [3] WU L, DENG S, XUAN J. Sampling and Pricing Strategy Under Competition [J]. Omega, 2018, 80(192-208).
- [4] HSU C L, LIN J C C, CHIANG H S. The effects of blogger recommendations on customers' online shopping intentions [J]. Internet Research, 2013, 23(1): 69-88.
- [5] BISWAS D, GREWAL D, ROGGEVEEN A. How the Order of Sampled Experiential Products Affects Choice [J]. Journal of Marketing Research, 2010, 47(3): 508-19.
- [6] LIN Z, ZHANG Y, TAN Y. An Empirical Study of Free Product Sampling and Rating Bias [J]. Information Systems Research, 2019, 30(1): 260-75.
- [7] FEHR E, XE, CHTER S. Fairness and Retaliation: The Economics of Reciprocity [J]. The Journal of Economic Perspectives, 2000, 14(3): 159-81.
- [8] SCHMIDHUBER, JÜRGEN. Deep learning in neural networks: An overview [J]. Neural Netw, 61(85-117).
- [9] NI F T, ZHANG J, NOORI M N. Deep learning for data anomaly detection and data compression of a long-span suspension bridge [J]. Computer-Aided Civil and Infrastructure Engineering,
- [10] XU W H, HUANG H, ZHANG J, et al. CNN-based Skip-Gram Method for Improving Classification Accuracy of Chinese Text [J]. Ksii Transactions on Internet and Information Systems, 2019, 13(12): 6080-96.
- [11] YADAV A, VISHWAKARMA D K. Sentiment analysis using deep learning architectures: a review [J]. Artificial Intelligence Review,
- [12] YANG W M, ZHANG X C, TIAN Y P, et al. Deep Learning for Single Image Super-Resolution: A Brief Review [J]. Ieee Transactions on Multimedia, 2019, 21(12): 3106-21.

- [13] DHILLON A, VERMA G K. Convolutional neural network: a review of models, methodologies and applications to object detection [J]. Progress in Artificial Intelligence,
- [14] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2014,
- [15] LAKSHMANAN A, KRISHNAN H S. The A-ha! Experience: Insight and Discontinuous Learning in Product Usage [J]. Journal of Marketing, 2011, 75(105-23.
- [16] KIM J, MORRIS J D. The Power Of Affective Response And Cognitive Structure In Product-Trial Attitude Formation [J]. Journal of Advertising, 2007, 36(1): 95-106.
- [17] CHEN H, DUAN W, ZHOU W. The interplay between free sampling and word of mouth in the online software market [J]. Decision Support Systems, 2017, 95(82-90.
- [18] LI F, YI Z. Trial or No Trial: Supplying Costly Signals to Improve Profits\*: Trial or No Trial: Supplying Costly Signals to Improve Profits [J]. Decision Sciences, 2016, 48(4): 795-827.
- [19] HAN Y, ZHANG Z. Impact of free sampling on product diffusion based on Bass model [J]. Electronic Commerce Research, 2018, 18(1): 125-41.
- [20] CONNELLY B, CERTO T, IRELAND R, et al. Signaling Theory: A Review and Assessment [J]. Journal of Management - J MANAGE, 2011, 37(1): 39-67.
- [21] KIRMANI A, RAO A. No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality [J]. Journal of Marketing, 2000, 64(2): 66-79.
- [22] HAMPSHIRE K, HAMILL H, MARIWAH S, et al. The application of Signalling Theory to health-related trust problems: The example of herbal clinics in Ghana and Tanzania [J]. Social Science & Medicine, 2017, 188(109-18.
- [23] MAVLANOVA T, BENBUNAN-FICH R, KOUFARIS M. Signaling theory and information asymmetry in online commerce [J]. Information & Management, 2012, 49(5): 240-7.
- [24] BISWAS D, BISWAS A. The diagnostic role of signals in the context of perceived risks in online shopping: Do signals matter more on the Web? [J]. Journal of Interactive Marketing, 2004, 18(3): 30-45.
- [25] SPENCE M. Signaling in Retrospect and the Informational Structure of Markets [J]. American Economic Review, 2002, 92(3): 434-59.
- [26] DONATH J. Signals in Social Supernet [J]. J Computer-Mediated Communication, 2007, 13(1): 231-51.
- [27] JIANG Z, BENBASAT I. Virtual Product Experience: Effects of Visual and Functional Control of Products on Perceived Diagnosticity and Flow in Electronic Shopping [J]. J of Management Information



Systems, 2005, 21(111-48).

[28] TENG L, YE N, YU Y, et al. Effects of culturally verbal and visual congruency/incongruency across cultures in a competitive advertising context [J]. Journal of Business Research, 2014, 67(3): 288–94.

[29] EVANS J S B T. In two minds: dual-process accounts of reasoning [J]. Trends in Cognitive Sciences, 2003, 7(10): 454-9.

[30] YOO J, KIM M. The effects of online product presentation on consumer responses: A mental imagery perspective [J]. Journal of Business Research, 2014, 67(11): 2464–72.

[31] MACINNIS D, PRICE L. The Role Of Imagery in Information Processing: Review and Extensions [J]. Journal of Consumer Research, 1987, 13(4): 473-91.

[32] WANG Q, CUI X, HUANG L, et al. Seller reputation or product presentation? An empirical investigation from cue utilization perspective [J]. International Journal of Information Management, 2016, 36(3): 271-83.

[33] ZHANG S, LEE D, VIR SINGH P, et al. How Much Is an Image Worth? Airbnb Property Demand Estimation Leveraging Large Scale Image Analytics [J]. SSRN Electronic Journal, 2017,

**基金项目：** 本文系国家自然科学基金项目“信息不对称驱动的共享经济去中心化机制与风险的复杂性研究”（项目编号：71874131）的研究成果之一。

#### **作者贡献声明：**

蔡婧璇：采集、清洗和分析数据；论文起草；

吴江：提出研究思路；论文修订；

王诚坤：设计研究方案；进行实验。

#### **利益冲突声明：**

所有作者声明不存在利益冲突关系。

#### **支撑数据：**

支撑数据由作者自存储，E-mail: jingxuancai@whu.edu.cn。

[1] 蔡婧璇. dmdata.xlsx. Dealmoon 网站众测报告数据.