Classification of Online Customer Reviews for Digital Product
Innovation: A Motivation Perspective

Chenxi Zhang
Bingqing Xiong
Eric Lim
Chee-Wee Tan
Zeshui Xu

Follow this and additional works at: https://aisel.aisnet.org/sighci2021
Classification of Online Consumer Reviews for Digital Product Innovation: A Motivation Perspective

Chenxi Zhang  
Sichuan University  
zcx950731@163.com

Bingqing Xiong  
University of Science and Technology of China  
City University of Hong Kong  
bxiong2-c@my.cityu.edu.hk

Eric T. K. Lim  
University of New South Wales  
et.lim@unsw.edu.au

Chee-Wee Tan  
Copenhagen Business School  
ct.digi@cbs.dk

Zeshui Xu  
Sichuan University  
xuzeshui@263.net

ABSTRACT
With rapid technological advances, digital products are becoming increasingly prevalent. Although past studies have examined the contribution of online reviews extensively in the context of physical products, there is limited understanding of the contribution of online reviews in digital product innovation. To this end, this study reviews previous work related to online reviews of physical and digital products in an attempt to reclassify online reviews of digital products from the perspective of consumer motivation. Taking game-related app reviews as an example, we employed a topic modeling model to extract insights related to consumer motivation. An in-depth appreciation of consumers’ motivation from analyzing online reviews can yield invaluable insights in driving digital product innovation.

Keywords  
Digital Products, Online Reviews, Consumer Motivation, Digital Innovation, Classification

INTRODUCTION
Technological advances have altered the way consumers interact with product developers (Gonçalves et al., 2018). Consumers can post online reviews to share their appraisals and concerns about products and services with developers (Li et al., 2014). This is especially true for digital products. Because digital products can be updated over time through versioning, it gives rise to a relentless need for continuous innovation (Wiesbök et al., 2020) with consumers playing the dual role of co-innovators in such innovation processes (Ghose et al., 2011). Not only can consumers influence others’ receptivity of digital products via sharing their usage experience, but they can also contribute innovative ideas to steer the development of novel features in successive versions. Consequently, it is imperative for developers to gain an intricate understanding of the motivational forces driving consumers to offer disparate forms of feedback to enact appropriate strategies to engage the latter in digital product innovation.

Due to versioning in digital products, consumers, as an integral part of product innovation, possess a greater incentive to share product-related information in online reviews. Like physical products, consumers have an identical need for self-expression after experiencing digital products, be it positive or negative. But at the same time, distinct from physical products, consumers are also willing to provide recommendations for digital product improvements in online reviews because they are likely to have a better product experience if their recommendations were to be incorporated into subsequent versions. In other words, consumers, by identifying product deficiencies and offering ideas about product innovations, can hope for better usage experience in the future. For this reason, online reviews of digital products contain supplementary information in the likes of bugs and feature requests (Jha & Mahmoud, 2019) that can support developers in product innovation (Timoshenko & Hauser, 2019). Exploring online reviews from the standpoint of consumer motivation can hence deliver purposeful and targeted insights for digital product innovation (Gonçalves et al., 2018).

Indeed, a review of extant literature points to a dearth of research on the role of online reviews of digital products from the standpoint of consumer motivation. First, most studies on online reviews are primarily concentrated on physical products (Malik & Hussain, 2018; Wang et al., 2019; Jang & Seongsoo, 2019; Mitra & Jenamani, 2020), which in turn led to scholarly calls for a deeper understanding of the contribution of online reviews in digital product innovation (Wiesbök et al., 2020) due to fundamental distinctions between digital and physical products. Second, extant literature on the research and development of traditional products affirms the importance of consumers’ active involvement in new product development (Chang & Taylor, 2016). However, most consumers participated in the product development process through focus groups and surveys arranged by the firms, which, despite incurring massive amounts of time and effort, reach only a tiny fraction of consumers (Nambisan, 2002). In contrast, online reviews constitute an accessible, inexpensive, and spontaneous means of soliciting consumer input to drive product innovation (Qi et al., 2021).
Meanwhile, existing studies have examined the factors that influence consumers’ willingness to post online reviews but paid little attention to why consumers mention different aspects of content in their online reviews (Yoo et al., 2019). Due to the impact of online reviews on the performance of digital products (He et al., 2020), developers should understand the motivation of consumers to post different aspects of online reviews to develop appropriate strategies to get better online reviews and lead to more profits (Gonçalves et al., 2018). Table 1 lists selected literature on three types of digital product research based on online reviews.

After reviewing the existing research, we find that there are three main motivations for consumers to post online reviews of digital products, namely, self-expression, functional benefits, and reciprocity. Self-expression refers to how consumers share their thoughts, attitudes, and activities through posing online reviews (Flecha-Ortíz et al., 2021). Compared to physical products, consumers have a similar need for self-expression after experiencing digital products (Hennig-Thurau et al., 2004). Unlike physical products, consumers of digital products have more substantial incentives for reciprocity and functional benefits. Reciprocity is conceived as a benefit for individuals to participate in social exchange (Cheung & Lee, 2012). When consumers do not know each other, the kind of reciprocity that is relevant is called ‘generalized’ exchange, and the person who offers help to others is expecting returns in the future (Wasko et al., 2000). Functional benefits are the extent to which consumers’ desire to influence developers and improve using experience about digital products (Jang & Seongsoo, 2019). In the setting of constant iteration and dynamic interaction of digital products, consumers will be driven by the two motivations and post some different reviews from those of physical products.
reviews are divided into different words. Third, we for word segmentation. After word segmentation, filtered reviews for subsequent processing. Second, we use Jieba five words or less. We finally obtained 110,000 online reviews and short texts of removed some uninformative reviews and short texts of online complaints. We employed Latent Dirichlet Allocation (LDA) for topic modeling to delve into the different expressions of consumers in online reviews. LDA can be used to understand the topics distribution of each online review and the word distribution of each topic. The process of topic modeling mainly includes four steps: preprocessing reviews, determining the optimal number of topics, topic extraction, topic adjustment.

Effective preprocessing of reviews before modeling is a critical step in topic modeling. First, we filtered the initial reviews according to the following two ways. (1) We deleted duplicate reviews and non-Chinese reviews. (2) We removed some uninformative reviews and short texts of five words or less. We finally obtained 110,000 online reviews for subsequent processing. Second, we use Jieba for word segmentation. After word segmentation, filtered reviews are divided into different words. Third, we eliminate useless characters, punctuations, and meaningless interference words according to the stop words list. The stop words list used in this paper is established by adding meaningless high-frequency words based on the HIT stop words list.

The number of potential topics has a significant impact on the results of topic modeling. We determine the optimal number of topics by calculating the coherence value. This paper builds many LDA models with different topics and selects the optimal number of topics with the highest coherence value. After calculating the optimal number of topics, the processed online reviews are converted to calculate the frequency of the word and create a word-review matrix. The distribution of topics in the online reviews and the distribution of words in each topic can be obtained by the word-review matrix.

The analysis results may contain similar topics or noisy words, so the final topics and related words should be adjusted manually. We manually adjust the obtained topics in LDA results based on the following two criteria. First, if the frequently occurring words in each topic are semantically similar, we merge these topics. Second, when it is impossible to determine the meaning of a topic based on the words in the topic, we will look at online reviews related to each topic. If these reviews express similar meanings, we will merge these topics.

### METHODOLOGY

**Data Description**

We collected online reviews from a popular Chinese mobile app store using crawlers developed with python. We crawled 1.15 million app reviews in 12 game-related categories, including strategy games, mobile games, puzzle games, role-playing games, and more. We randomly selected one-tenth of the reviews from each category, and ultimately selected 11,500 app reviews to extract insights related to consumer motivation.

**Topic Modeling**

We employ Latent Dirichlet Allocation (LDA) for topic modeling to delve into the different expressions of consumers in online reviews. LDA can be used to understand the topics distribution of each online review and the word distribution of each topic. The process of topic modeling mainly includes four steps: preprocessing reviews, determining the optimal number of topics, topic extraction, topic adjustment.

A table is provided with three types of digital product research based on online reviews.

**ANALYTICAL RESULTS**

By calculating the coherence value, we find that when the number of topics is set at 14, the information divergence between topics reaches the maximum. Therefore, the optimal number of topics is 14 for the data in this paper. After careful review, two similar topics were adjusted, and 12 topics are finally identified. The final 12 topics are summarized in Table 2. We divided the final extracted topics into three different categories from three motivations (i.e., self-expression, functional benefits, and reciprocity). The results verify our proposed classification framework for online reviews of digital products from the perspective of consumer motivation. The usage experience category includes descriptions and evaluations of features and functions in apps. The feature request category includes language request and update request. The bugs report category includes bugs and original.

<table>
<thead>
<tr>
<th>Category</th>
<th>Topic</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage Experience</td>
<td>Description</td>
<td>Features</td>
</tr>
<tr>
<td></td>
<td>Interface</td>
<td>pixel, post, interface, screen, pattern</td>
</tr>
<tr>
<td></td>
<td>Operation</td>
<td>operation, loading, download, on-line</td>
</tr>
<tr>
<td></td>
<td>Version</td>
<td>Android, Apple, system</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Feature</td>
<td>storyline, bgm, music, skill, image</td>
</tr>
<tr>
<td></td>
<td>Interface</td>
<td>frames, painting style, image quality</td>
</tr>
<tr>
<td></td>
<td>Difficulty</td>
<td>difficulty, simple, clear the game</td>
</tr>
<tr>
<td></td>
<td>Rating</td>
<td>like, good, boring, thumb-up</td>
</tr>
<tr>
<td>Feature Request</td>
<td>Language request</td>
<td>English, Chinese version, translate</td>
</tr>
<tr>
<td></td>
<td>Update request</td>
<td>update, upgrade, add, request, improve</td>
</tr>
</tbody>
</table>
CONCLUSION AND FURTHER RESEARCH

Given the different characteristics of traditional and digital products, this paper considers three consumer motivations for posting online reviews of digital products and reclassifies online reviews from these motivations to help fully understand the multiple aspects of online reviews. This study takes app reviews as an example and uses a topic modeling technique to identify insights related to consumers’ motivation in app reviews. An in-depth appreciation of consumers’ motivation by analyzing the extracted topics enables developers to better understand consumers’ demands and inform digital product innovation.

In the future, we will continue to explore three aspects based on the research basis of this paper. First, we will analyze what drives consumers to spread online reviews on digital platforms through online questionnaires. Second, we will further refine our classification of online reviews from the perspective of consumer motivation and supply recommendations to developers based on the impact of different types of online reviews on product innovation and performance. Third, in addition to providing suggestions for developers, we will also examine how to design different functions of digital platforms to motivate consumers to provide more online reviews and further drive digital innovation.

REFERENCES


Table 2. Samples of Extracted Topics and Keywords

<table>
<thead>
<tr>
<th>Bugs Report</th>
<th>Bug</th>
<th>crash, blank screen, bug, stutter, delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>plagiarize, imitation, cheat, copy, pirate</td>
<td></td>
</tr>
</tbody>
</table>

Zhang et al. Classification of Online Consumer Reviews for Digital Product Innovation