

Article

User Experience Quantification Model from Online User Reviews

Jamil Hussain ¹, Zahra Azhar ², Hafiz Farooq Ahmad ³, Muhammad Afzal ⁴, Mukhlis Raza ⁵
and Sungyoung Lee ^{6,*}

¹ Department of Data Science, Sejong University, Seoul 05006, Korea; jamil@sejong.ac.kr

² Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, CA 95064, USA; zazhar@ucsc.edu

³ Computer Science Department, College of Computer Sciences and Information Technology (CCSIT), King Faisal University, Al-Ahsa 31982, Saudi Arabia; hfahmad@kfu.edu.sa

⁴ Department of Software, Sejong University, Seoul 05006, Korea; mafzal@sejong.ac.kr

⁵ Department of Faculty of Computing, Riphah International University, Islamabad 46000, Pakistan; 2066@students.riphah.edu.pk

⁶ Department of Computer Science and Engineering, Kyung Hee University, Seocheon-dong, Giheung-gu, Yongin-si 17104, Korea

* Correspondence: sylee@oslab.khu.ac.kr

Abstract: Due to the advancement in information technology and the boom of micro-blogging platforms, a growing number of online reviews are posted daily on product distributed platforms in the form of spontaneous and insightful user feedback, and these can be used as a significant data source to understand user experience (UX) and satisfaction. However, despite the vast amount of online reviews, the existing literature focuses on online ratings and ignores the real textual context in reviews. We proposed a three-step UX quantification model from online reviews to understand customer satisfaction using the effect-based Kano model. First, the relevant online reviews are selected using various filter mechanisms. Second, UX dimensions (UXDs) are extracted using a proposed method called UX word embedding Latent Dirichlet allocation (UXWE-LDA) and sentiment orientation using a transformer-based pipeline. Then, the casual relationships are identified for the extracted UXDs. Third, the UXDs are mapped on the customer satisfaction model (effect-based Kano) to understand the user perspective about the system, product, or services. Finally, the different parts of the proposed quantification model are evaluated to examine the performance of this method. We present different results of the proposed method in terms of accuracy, topic coherence (TC), Topic-wise performance, and expert-based evaluation for the proposed framework validation. For review quality filters, we achieved 98.49% accuracy for the spam detection classifier and 95% accuracy for the relatedness detection classifier. The results show that the proposed method for the topic extractor module always gives a higher TC value than other models such as WE-LDA and LDA. Regarding topic-wise performance measures, UXWE-LDA achieves a 3% improvement on average compared to LDA due to the incorporation of semantic domain knowledge. We also compute the Jaccard coefficient similarity between the extracted dimensions using UXWE-LDA and UX experts-based analysis for checking the mutual agreement, which is 0.3, 0.5, and 0.4, respectively. Based on the Kano model, the presented study has potential implications concerning issues and knowing the product's strengths and weaknesses in product design.

Keywords: customer satisfaction; online reviews; Kano model; product improvement; sentiment analysis; opinion mining; user experience



Citation: Hussain, J.; Azhar, Z.; Ahmad, H.F.; Afzal, M.; Raza, M.; Lee, S. User Experience Quantification Model from Online User Reviews. *Appl. Sci.* **2022**, *12*, 6700. <https://doi.org/10.3390/app12136700>

Academic Editor: João M. F. Rodrigues

Received: 3 June 2022

Accepted: 30 June 2022

Published: 1 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Consumers of contemporary society desire innovative products that generate positive and initiative experiences. With this in mind, most product designers focus on the relationship between positive user experience (UX) and product design and success [1,2]. Various

factors contribute to establishing a positive UX (e.g., user satisfaction, context-of-use, quality, enjoyment, ease of use, and others). Therefore, thoroughly comprehending the UX of a target product, system, or service is essential to nurturing consumer relations. Although several studies have discussed various methods for assessing and evaluating UX, no one method has been universally accepted as UX is context-dependent, subjective in nature, and quite dynamic.

Furthermore, UX is broadly described as consisting of user sentiments regarding a product, system, or service [3,4]. According to ISO 9241-11:2018(E) [5], UX is described as: “person’s perceptions and responses resulting from the use and anticipated use of a product, system or service.” UX influences factors such as the user’s mental and physical state, product, and contexts of use that occur before, during, and after use [6]. Additionally, many studies have asserted that a positive UX plays a vital role in motivating user loyalty, such as recommendations of products to family, positive reviews, or continuous usage.

Most of these prior studies employed traditional methods such as questionnaires, surveys, report grand techniques (RGT) in the field, and lab studies to evaluate UX by crafting various scenarios [7–9]. In these scenarios, the UX moderator defined various tasks and context-of-uses during the participants’ interaction with the products [3]. Additional parts of this method include task arrangement, participant selection, UX evaluation methods and training, and cost involving collecting sample data. Although these methods are crucial to collecting essential user experience data, such approaches consider limited aspects of data collection and may lead to a more significant impact on product-related sentiments. Moreover, the measurement items used in the surveys in prior studies were developed based on possibly inconsistent knowledge and disregard for end-users perspectives.

Furthermore, existing literature has revealed user reviews to extract valuable information about consumer preferences experienced during product usage. Additionally, user reviews are obtained from a diverse sample. Different users associate with various performances and experiences for the same products. Such data give rise to a more thorough understanding, bettering new product designs.

Various approaches have been developed to obtain the different insights from the online user review. Despite the vast amount of online user reviews, some existing literature primarily focuses on online numerical ratings, ignoring the actual textual context in online user reviews. The textual context often contains relevant and profitable information such as features requested and bug reports which can significantly aid product advancements. However, this vast number of user reviews in the unstructured form is written in natural language. Being able to process reviews that allow for developing new products and improving existing ones is a currently unfulfilled necessity. There is also a lack of a method that extracts UX information from user reviews with embedded requested features. Furthermore, applying text mining techniques to derive UX insights from extensive UGC data is quite challenging.

Sentiment analysis and opinion mining are often used to find users’ opinions of a product [10], but the extraction of UX information from user reviews is limited nonetheless [11]. Correspondingly, evolving research in the sphere of UX studies entails various attempts to investigate consumer experiences from online user reviews. These studies can be classified into two categories: (1) mining the user experience aspects or dimensions (UXDs) from online reviews [12] and (2) modeling UX from online user reviews [13].

In the first category, numerous text mining and machine learning techniques are employed for the extractions of different UX aspects, such as probabilistic topic models: Latent Dirichlet Allocation (LDA), Probabilistic Latent Semantic Analysis (PLSA) [12,14,15], word embedding, aspect-based sentiment analysis [16], and analyzing the relative importance of each extracted UX aspect. In the second category, researchers try to develop a mapping mechanism of all their target extracted UX dimensions on the existing user satisfaction models, such as the Kano model, to give a road map for a product, system, or service improvement or development.

We designed a comprehensive framework for modeling UX from online reviews to resolve these challenges. First, we filtered user reviews unrelated to the UX domain using UX multi-criteria qualifiers. Then, we extracted UX aspects from the filtered user reviews using an enhanced topic extraction methodology called UXWE-LDA. UXWE-LDA improves existing knowledge-based topic models by extracting more domain-dependent dimensions in the UX area through UGC. It combines topic modeling, specifically LDA, with word embedding that automatically learns the domain knowledge from a large amount of textual data. The proposed method gains domain knowledge from the vast number of documents using co-occurrence and word-embedding word vectors correlation of related data, resulting in a more coherent topic. Then, sentiment analysis is applied to reviews concerning the extracted UX aspects or dimensions.

We aimed to extract the essential aspects by inducing a positive UX that utilized UGC data. Mining UXDs from UGC data allows us to comprehend customer preferences and needs effectively and reliably, allowing the product owner to improve their product design, system, or service. The presented study has potential implications for product design. It can mine the most concerning UX aspects from online reviews, allowing the withdrawal of valuable information for effective product redesign. Furthermore, it can identify the strengths and weaknesses of a product according to the Kano model. This method allows the product designer to understand the different categories of UXDs in the UEQ model, therefore establishing its crucial role in product enhancement. According to the classification results of UXDs, the priority order of UXDs enables developers to plan product enhancements. More specifically, the contributions are made in three parts.

- First, the user quality filter module identifies user reviews containing helpful information related to UX. This step is essential to removing trivial user reviews before applying topic modeling. This module classifies online reviews based on predefined UX aspects (user, situation, and product facets).
- UXDs extraction from online reviews using proposed user experience word embedding LDA (UXWE-LDA) methodology allows for the automatic learning of the domain knowledge from the given text corpus to generate a more coherent topic. It mainly contains two steps: UXWE-LDA and sentiment analysis. The UXWE-LDA is an improved version of LDA that takes the domain knowledge from the given text corpus, extracts more coherent topics, and assigns labels as UXD to each extracted topic using a dictionary-based approach. Then, it identifies the sentiment orientations of the reviews concerning each UXD based on ensemble methodology. Finally, it classifies each review into positive or negative sentiment categories and associates the sentiment orientation with the extracted UXDs.
- The causal relationship of sentiments toward each UXD on user satisfaction obtained from using the Bi-LSTM model to overcome the problem of existing models of review-based user satisfaction studies.

The rest of the paper is structured as follows. Section 2 discusses the materials and methods for the analysis conducted. Section 3 describes the results and case study based on the proposed methodology, Section 4 presents a discussion, and Section 5 concludes the work.

2. Related Work

The key focus of this research is to understand the current research work that maps dimensions to aspects, phenomena, and viewpoints in UX. A brief description of those research works is as follows.

2.1. Dimensions of Usability and UX

The usability defined by the ISO 9241 standard [5] uses three dimensions: efficiency, effectiveness, and satisfaction. They define usability as “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use”. A detailed description of usability is mapped

to five dimensions by [17,18]. However, in the literature, there are some deviations and variations in the naming of dimensions [19]. So, the final five dimensions, we are focusing on include: (i) Effectiveness/Errors, (ii) Efficiency, (iii) Satisfaction, (iv) Learnability, and (v) Memorability.

Compared to usability, there exists a minimal consensus on UX definition and its mapping to aspects. Researchers defined and characterized UX from many perspectives to link it to their academic and application aims. While some researchers believe that UX is holistic, others claim that complex experiences, such as emotions and usability, are generated by summative and evaluative constructs [20].

Furthermore, some researchers highlight the significance of UX characteristics such as product features, user state, and contexts [21], while others have found associations between usability and UX for heterogeneous factors such as gender differences, the context of use, and usage patterns [22]. When the emphasis is on certain UX aspects and their connections, user input is often gathered and evaluated based on these factors. According to the ISO 9241-210 [5], UX is defined as: “A person’s perceptions and responses that result from the use and anticipated use of a product, system or service”.

So, a UX dimension is a key or essential component that may explain how a UX is created. Based on previous UX research, UX is defined by a system’s pragmatic (“instrumental product”, “task-oriented”, or “ergonomic”) and hedonic (“non-instrumental”, “non-task-oriented”) qualities dimensions [23,24]. Pragmatic quality is the degree of usefulness, efficiency, and simplicity of usage. Hedonic traits include “joy of use”, focus evoking, identification, and stimulation. Hedonic quality is the total of pragmatic attributes that might trigger positive or negative emotions and affect a product’s acceptance [23].

2.2. Usability and UX in Online User Reviews

Usability measures the overall ability of a product, service, or system to achieve targeted goals effectively and proficiently, while UX evaluations provide a perception of the users’ satisfaction towards achieving these goals. Both usability and UX are closely related to the specific product, defined task, user cognitive, and distinct circumstances. They play an essential role in critical product analysis and are the target of academic evaluations. Product reviews are the rich sources of identifying the usability and UX of a targeted product. It helps in understanding user opinion about a product and assists in product improvements. Potential users typically check the reviews given by other users to make a final decision of whether to purchase a product. Additionally, the reviews reveal the real UX of a user about the product as it is given after consuming the services and using the product. The user provides product feedback in the form of reviews due to motivation, tangible, and intangible rewards.

Despite benefits, there are some limitations to considering online product reviews for usability and UX evaluation. The reviews strongly describe user opinion towards a product. However, in user online reviews, some important information is missing such as age, gender, and preferences, which are required for usability studies. Moreover, all reviews are not credible for usability study; some reviews may contain false information or are even provided by the owner of the product to promote their products.

2.3. Mining the UX Dimensions from Online Reviews

An evolving stream of UX research has focused on assessing the UX directly or indirectly from online reviews. Online user reviews are real reservoirs of the UX. These are unstructured textual documents containing a large amount of information. The quantitative analysis of these reviews generates insight by applying text mining and analytics techniques. Additionally, these techniques extract important information from unstructured text data and then analyze such information. Currently, text mining is intensifying the major research areas of sentiment analysis, topic modeling, document classification, and natural language processing. Generally, the studies in these domains can be categorized into extracting UXDs from online reviews and modeling UX from online reviews [14].

The user experience dimensions mining extracts the UXDs from online user reviews and evaluates the equal importance of each UXD. For instance, Tirunillai and Tellis [25] proposed a framework for extracting UX aspects from online reviews through an improved LDA model. Yue Guo (2017) [12] used data from 266,544 online reviews, topic modeling, and content analysis to analyze user satisfaction. Likewise, nearly all similar studies [12,14,15,26] engaged in the topic modeling approach, specifically LDA, for the extraction of latent dimensions and conduction of a regression analysis focusing on the rating data for the verification and validation of the extracted dimensions or aspects in the domain of UX.

Several studies categorize online user reviews by applying sentiment analysis as positive, negative, or neutral. Suryadi et al. [16] used NLP and machine learning techniques to identify aspect-based sentiment for various components in particular contexts. Such analysis allows observing a product's status against its competitors in a specific context. Combining online ratings and content analysis of reviews by NLP and machine learning enables researchers to identify the causal relationship between extracted UX aspects and consumer satisfaction. Yang et al. [11] presented a machine-learning-based technique to assess the user's UX using online customer reviews. This technique provides UX assistance for product design optimization and supports UX research.

Currently, researchers often attempt to apply other word representation schemes such as word embedding into topic modeling, reducing the dimensionality of word vectors based on the co-occurrence information by considering the local context of words and combining the global and local context to provide more cohesive topics. However, the unsupervised models frequently generate semantically incoherent topics that are difficult to understand [27,28]. Some previous works add domain knowledge in the topic modeling to resolve the shortcomings of unsupervised models, but most models cannot learn domain knowledge automatically [29].

2.4. Modeling UX from Online Reviews

Various studies have been proposed to model UX and user satisfaction from online reviews in the second category. The modeling UX from online reviews primarily examines the effects of user sentiments towards product features on UX, particularly on customer satisfaction.

Farhad et al. [13] proposed a Bayesian approach using semi-structured data for aspect-level sentiment analysis and UX modeling. They associated the sentiment with the product aspect in each review using a probabilistic approach to producing a single rating for each attribute and their relative importance to the product or service.

Similarly, Decker et al. [30] used regression models (Poisson, negative binomial, and latent class Poisson) to assess user sentiments' effects on product aspects on user satisfaction. Their results reveal a negative binomial regression model to outperform similar models in identifying the causal impact of user sentiments towards product aspects on user satisfaction.

While these studies have made substantial contributions to modeling UX and review-based user satisfaction investigations, they entail complex components such as their reliance on the supposition that the online rating follows an unstable Gaussian distribution. Additionally, the Kano model developed by Kano et al. [31] was used in existing studies for modeling customer satisfaction. This model categorizes the product features in classes such as must-be, performance, excitement, indifferent, and reverse. These feature values associate with user satisfaction [14].

We propose a new method for evaluating online consumer reviews for UX modeling. Due to the lack of research on how to use UX analysis to improve product design, this article focuses on using the User Experience Questionnaire (UEQ) to combine hedonic and pragmatic qualities into UX modeling. The proposed method may reduce the UX research gap by accelerating UX exploration and optimizing product and service experiences.

3. Materials and Methods

We proposed the three-step methodology for modeling user satisfaction from online user reviews, shown in Figure 1.

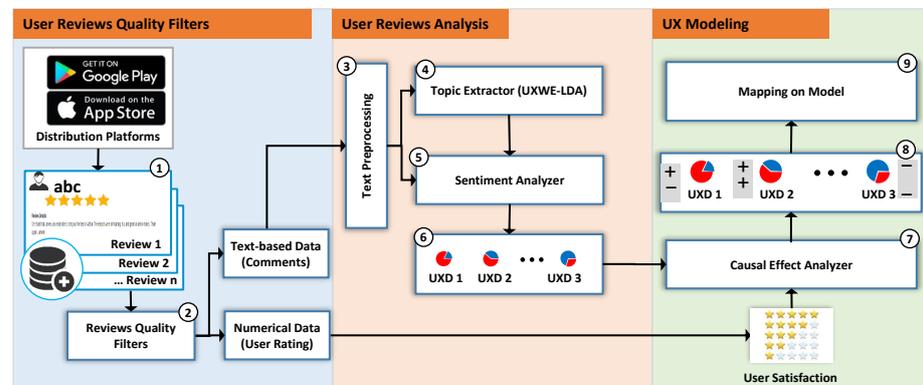


Figure 1. Abstract view of proposed methodology.

First, the usefulness of online user reviews containing information related to UX and usability from the corpus collection will be identified. Before applying the user review analysis, the framework checks the quality and reliability of user reviews through three user review quality filters: spam detection, relatedness, and subjectivity of review documents. These three classifiers function in a sequential format. They filter spam reviews, access UX-related reviews, and then select the subjective reviews. The relatedness classifier, also known as the UX multi-criteria qualifier (UXMCQ), uses a mainly unsupervised method requiring minimal configuration of domain seed words to auto-label the data based on the context window (see Section 3.1 for more details).

The second step process consists of the following: (i) UX dimensions (UXDs) extraction using the proposed user experience word-embedding LDA (UXWE-LDA), an improved knowledge-based topic modeling methodology, and (ii) sentiment analysis and its orientation for each extracted UXD from online user reviews. UXWE-LDA is an improved LDA version that automatically learns the domain knowledge from the given text corpus. UXWE-LDA resolves the problems of existing LDA, often generating semantically incoherent topics. UXWE-LDA improves the existing knowledge-based topic models by UGC extracting more domain-dependent dimensions in the UX area. UXWE-LDA combines topic modeling, specifically LDA, with word embedding that automatically learns the domain knowledge from a large amount of textual data. This model automatically learns the domain knowledge from the given text corpus and extracts more coherent topics to assign labels as UXD to each extracted topic using a dictionary-based approach. Additionally, it identifies the user's positive and negative sentiment association towards each UXD. To identify the sentiment orientation, we employed the BERT-based sentiment transformer pipeline.

The third step consists of two parts: (i) casual relation analysis of UXDs with respective sentiment orientation and (ii) mapping the UXDs' causal relationship on the user satisfaction model. This overcame the problem of existing models based on measuring satisfaction from online reviews. The Bi-LSTM model combines the user rating and extracted dimensions to measure user sentiment's causal relationship on user satisfaction. In addition, we employed the two-dimensional Kano model for user satisfaction. Developed by Kano et al. [31], this model categorizes the product features into different classes: must-be, performance, excitement, indifferent, and reverse. These features' values are associated with user satisfaction [14]. The subsequent section describes each step in greater detail.

3.1. User Review Quality Filters

Before applying the user review analysis, the framework checks the quality and reliability of user reviews through user review quality filters such as spam detection,

relatedness, and subjectivity of online review documents. We employed three classifiers for a quality and reliability check to boost the topic coherence in the topic extraction process. These classifiers function in the following sequence: filter spam reviews, check for UX-related reviews and select the subjective reviews for UX modeling, as shown in Figure 2.

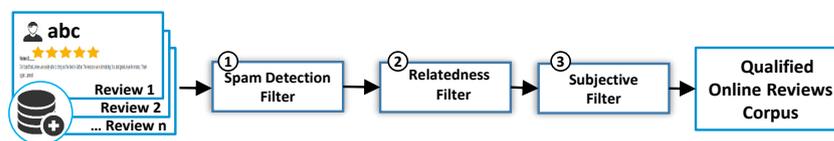


Figure 2. Abstract view of user review quality checking process.

3.1.1. Spam Detection Classifier

The spam detection classifier confirmed the quality of online user reviews to be either truthful or deceptive. Unfortunately, product distribution platforms, such as the Google Play store and Apple store, are frequently abused as potentially malicious users can freely insert fraudulent information without validation. Consequently, online review systems can become targets of individual and professional spammers, who insert deceptive reviews by manipulating the reviews’ ratings and content. When training the spam detection classifier, we used the “Deceptive Opinion Spam Corpus v1.4” [32] as the training dataset. In addition, the “ktrain” Python library [33] was used for training the spam detection 235 classifiers.

3.1.2. Relatedness Detection Classifier

Before applying topic modeling, it is essential to filter out reviews that contain data unrelated to a specific domain; we proposed a primarily unsupervised ML approach called UX multi-criteria qualifier (UXMCQ) to detect the relatedness of review related to the UX domain. This type of filter can boost the topic coherence in the topic extraction methodology. Thus, the UXMCQ selects reviews containing helpful information relating to UX for topic modeling.

The UXMCQ model creation mainly consists of three steps: (i) UX aspects dictionary creation and aspects configuration; (ii) word occurrence mapping and context window creation for auto labeling; and (iii) model creation and training. The overall process model is shown in Figure 3. The details of each step are described in the following subsections.

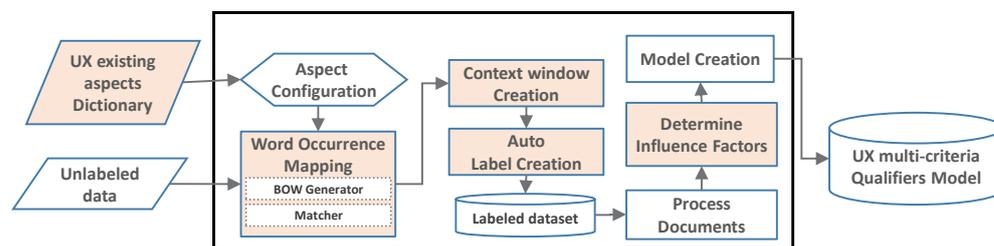


Figure 3. A UX multi-criteria qualifiers model overview.

UX Aspects Dictionary Creation and Aspects Configuration

UX aspects configuration is the primary step for the UXMCQ module. Based on the selected aspects, the model automatically labeled the unlabeled data using the bootstrap method based on the occurrence of a word using the context window size. It is essential to make the domain depend on aspect seeds for filtering the critical reviews for UXDs extraction. In order to make the UX domain aspects, we made the UX aspects dictionary using a systematic review process.

As mentioned earlier, UX is context-dependent, subjective in nature, and dynamic. As prior research has considered numerous aspects for measuring UX, scanning related studies is critical to designing a systematic review process that identifies the UX dimensions or aspects in the UX domain, allowing for the construction of a UX aspects dictionary for aspect configuration. This UX aspect dictionary can help build a more comprehensive UX model for UX evaluation. We used a two-phase approach for the extraction of UX aspects, as shown in Figure 4. In the first phase, we used the systematic review process to identify the UX-related literature which mentioned the UX aspects, dimensions, and measurements. In the later phase, we analyzed the selected papers for UX aspects selection. Finally, we constructed the UX aspects dictionary.

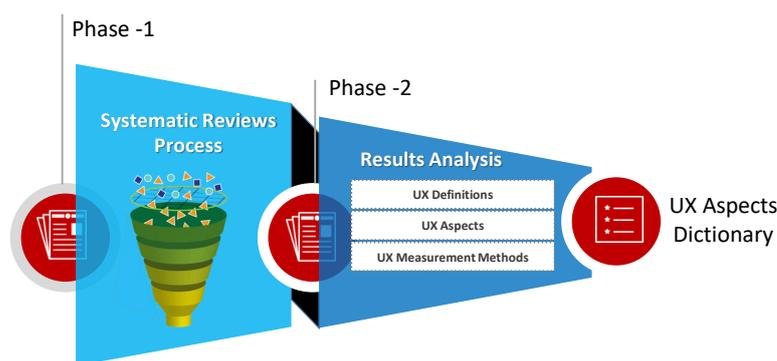


Figure 4. Identifying UX existing aspects for aspect configuration process.

A systematic review process was used for article selection in UX research. First, the publications were selected using four steps borrowed from [34]. We grouped the UX aspects based on the existing conceptual UX Facet model [11]. The UX facet model divided all essential factors into three main facets: user facet, product facet, and situation facet. The user facet is related to user sentiment and cognition, such as background information, user preferences, intentions, and opinions (negative, positive, or neutral). Product facet is related to product attributes such as UI, aesthetic, quality, and others. Finally, the situation facet is related to the environmental factors of the context of use, such as time and place.

For the third UX facet (situation facet), we used the Linguistic Inquiry and Word Count (LIWC) (<http://liwc.wpengine.com/> (accessed on 30 March 2022)) tool categories including “Time”, “Space”, and “Work”. Furthermore, as the LIWC tool reveals common thoughts, emotions, feelings, moods, personal and social concerns, and motivation, it was used to analyze the given text based on the dictionary. The percentage was calculated based on how well the words of the given text matched to the dictionary categories.

Aspect Configuration

UXMCQ only requires a small amount of domain aspects as seed words. As aforementioned, we created the UX aspects dictionary for aspect configuration. According to the context window, the aspects seed words are used as gold standards to auto-annotate the unlabeled data based on the occurrences of these seed words.

Word Occurrence Mapping and Context Window Creation for Auto Labeling

We used the bootstrap method for auto labeling based on the gold aspect terms related to the three UX facets. The auto labeling is based on the occurrence of the term by exact matching with the aspect terms in the unlabeled data. We used the context window of the size [+3, −3] and generated the label as UX facets based on matching aspect terms. The overall bootstrapping process is described in Algorithm 1.

Algorithm 1: Aspect based Auto Labeling

```

Input :  $D = \{d_1, d_2, d_3, \dots, d_n\}$  // Collection of user Reviews
          $A_T = \{T_1, T_2, T_3, \dots, T_k\}$  // Aspect Terms
Result: Labeled Data Based On Aspects
foreach  $d_i$  in  $D$  do
    foreach  $ConceptC$  in  $operands$  do
         $S = \{s_1, s_2, \dots, s_n\}$  // Split all reviews into sentences
        for  $s_i$  in  $S$  do
             $Matched \leftarrow Match(s_i, T)$  // Match the aspect Term in each Sentence S
            if  $Matched = True$  then
                 $C_w \leftarrow$  create context window  $[+n, -n]$ 
                 $Label \leftarrow$  Assign label to that context window as  $T$ 
            end
        end
    end
end

```

Figure 5 depicts how the bootstrap method assigns labels based on the terms occurrence and context window. First, it loads all the unlabeled review data, splits each review into sentences, then the matcher checks the occurrence of the aspect terms in each review; if a match is found it creates a context window and assigns labels as aspect terms.

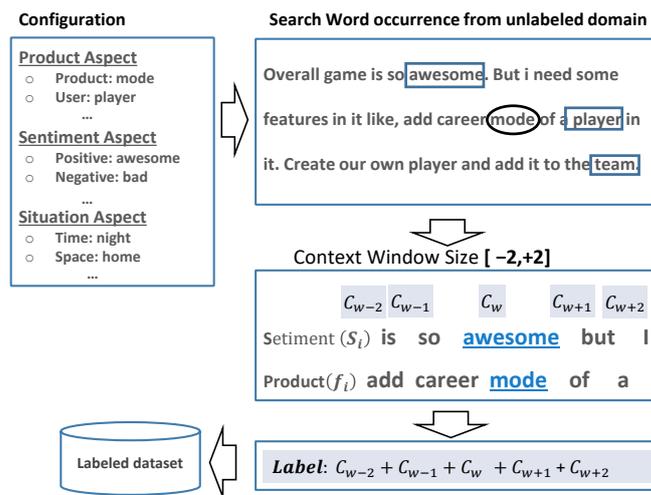


Figure 5. Auto labeling process based on the context window.

Model creation and training

We have employed the BERT-based model for training the UXMCQ classifier using “ktrain” Python library [33]. This model classifies the user reviews into either UX qualifiers or none.

3.1.3. Subjective Filter

Subjectivity identification is a key aspect of a person’s opinion, and we can classify online reviews as opinionated or not opinionated. We employed the existing Python library called TextBlob [35] for this task, which gives subjectivity/objectivity classification in the range [0.0, 1.0] where 0.0 is a very objective sentence, and 1.0 is very subjective.

3.2. User Review Analysis

In the user review analysis module, we presented the process of (i) UX dimensions (UXDs) extraction using the proposed user experience word-embedding LDA (UXWE-LDA)

topic modeling and (ii) sentiment analysis and its orientation for each extracted UXD from online user reviews. UX Dimensions (UXDs) extraction

UX Dimensions Extraction

For UXDs extraction, we developed a User Experience Word-Embedding LDA (UXWE-LDA) model that can extract more coherent topics by learning domain knowledge automatically from a given text corpus and assigning labels as UXDs using the dictionary-based approach. UXWE-LDA improves the existing knowledge-based topic models by extracting more domain-dependent dimensions in the UX area through UGC. UXWE-LDA combines topic modeling with a word embedding approach that automatically learns the domain knowledge from a large amount of textual data. UXWE-LDA workflow mainly consists of four steps, as shown in Figure 6. A detailed description of this model is given in the following sections.

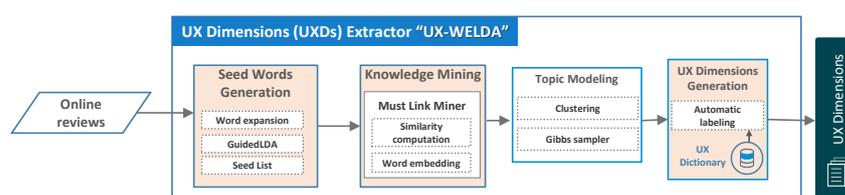


Figure 6. Abstract workflow for UX Dimensions Extractor.

Seed Words Generation

This step generated the global context from collections of online review corpora. First, all reviews are processed to convert the unstructured text into a structured form. For preprocessing, we applied tokenization, stemming, filter stop words, and others. For seed word generation, we used two-step processes. First, we ran a guidedLDA with guided seed words and selected topical words as seed words. We used the same methodology as [29] for seed word generation but, internally, our method’s syntactic and semantic relationships were unique. We used the guided LDA instead of a simple LDA to generate the seed words of interest. Second, we expanded the produced seed words using pre-trained word embedding models to make a more comprehensive global context. Algorithm 2 explains the overall process.

We enhanced the global seeds generated by the guided LDA. This considers the syntactic variation of the words (w) and the semantic similarity of a given corpus. In the existing literature, semantic similarity is computed using a manually built dictionary [36]. The issues with dictionary approaches include extensive human involvement, effort, and time required to hand-craft the dictionary. It is also challenging to scale a dictionary to incorporate the new contexts. Currently, researchers are attempting intuitive ways to compute semantic similarities using word distances, but they often disregard the context of the words in word embedding spaces [37]. Most of the prior works only focus on the implicit relationship in a word context window within the document [38], but do not consider the similarity of the word with pre-trained word embedding models. We used a similar approach called CluWord [39] to exploit the word similarity based on a pre-trained word embedding model to create a more general global context in semantic and syntactic terms. We used the Word2Vec [40] for pre-trained word representation using googleNews data. Let G_V represent the global vocabulary generated by guideLDA for all documents topics D_T . Let W_E be the word embedding vector representation for each term in G_V based on the pre-trained word embedding model. We compute the word expansion based on the following Equation (1). The Table 1 shows an example of word expansion for the word “chat”.

$$W_{t,t'} = \begin{cases} \delta(t, t') & \text{if } \delta(t, t') \geq \alpha \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $\delta(t, t')$ is computed using cosine similarity, matching the definition in Equation (2), and α is the threshold value for filtering the most similar words to t .

$$\delta(t, t') = \frac{\sum_{i=1}^n u_i v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \quad (2)$$

Regarding the δ_t for term t , the expansion is limited based on the α value to remove the unrelated words that have no significant relationship to term t . If the similarity between t and t' is less than the threshold value, we discard the t' .

Algorithm 2: Seed words generation algorithm

Input : Useful reviews corpus C ,
 Seed topic words S_d
 External corpus C
 Vector dimension k
 Vector dimension k

Result: The global context for user reviews text W_t

foreach (*document* $d \in C$) **do**
 | Sampling a topic form a topic's multiple distribution.
 | $Z_d \sim \text{Mul}(\theta)$
 | **foreach** *word* \in *document* d where $W \in (wd_1, wd_2, \dots, wd_n)$ **do**
 | | Generate a variable weight probability from the Bernulli distribution the
 | | probability of under t estimated by guided LDA
 | | $W_t = n - \text{argmar}_w \mathcal{O}(w, s_d)$
 | **end**
end

$W2V\text{Train}(C, k)$
 $\text{VocabSize} \leftarrow \text{getVocabSize}(C)$
 $V \leftarrow \text{initVector}(\text{vocabSize}, k)$
 $\theta \leftarrow \text{initVector}(\text{vocabSize}, k)$

for ($W_i \in C$) **do**
 | $e \leftarrow 0$
 | $Xw \leftarrow \sum u \in \text{context}(W_i) V(u)$
 | **for** ($u = w_i \text{UNEG}(w_i)$) **do**
 | | $e \leftarrow e + g\theta^u$
 | **end**
 | **for** ($u \in \text{Context}(w_i)$) **do**
 | | $V(u) \leftarrow -V(u) + e$
 | **end**
end

$t' = V(W_t, k)$
 $\text{expendedWord} \leftarrow 0$
for ($t' \in W_t$) **do**
 | $\delta(t, t') = \frac{\sum_i^l u_i v_i}{\sqrt{\sum_i^l u_i^2} \cdot \sqrt{\sum_i^l v_i^2}}$
 | **if** ($\delta(t, t') > \alpha$) **then**
 | | $\text{expendedWord} \leftarrow t'$
 | **end**
end

$W_t = W_t + \text{expendedWord}$
 return W_t

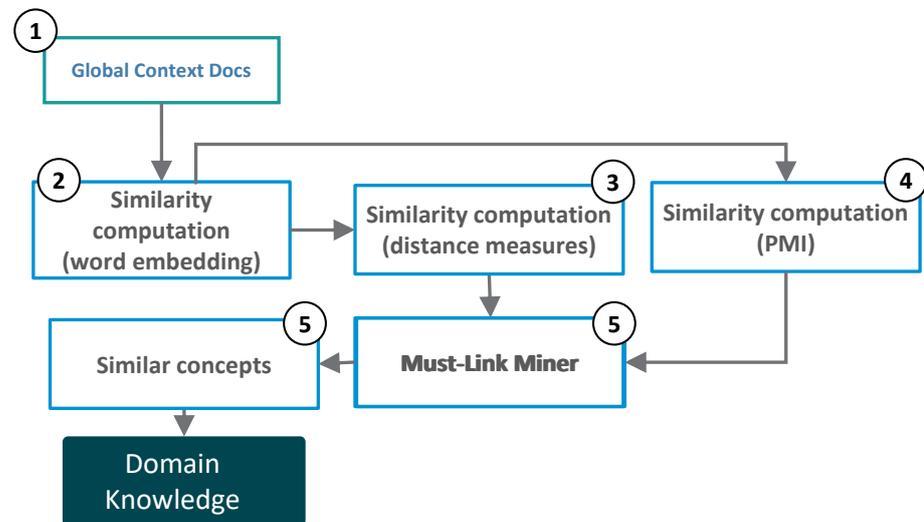
Table 1. Word expansion example based on the pre-trained word embedding model.

Similarity	Term t: chat
Semantically	chatroom, conversation , conversing, talk, conversed, Live_Chat, message, interview, speak,
Syntactically	Chats, chatting, Chat, chatted

Finally, we created the global context document (G_d) for each expended topic.

Knowledge Mining

We incorporated word-embedding and two other similarity computations for knowledge mining, including concise similarity and PMI. The overall process of knowledge mining is shown in Figure 7.

**Figure 7.** Workflow of must-link mining using similarity computation.

We computed the similarity between each word using cosine similarity based on the generated vectors trained by the Word2Vec. The cosine similarity of word vectors \vec{V} and \vec{U} is computed for w_1 and w_2 using Equation (3).

$$\text{sim}(w_1, w_2) = \frac{\vec{U}_{w_1} \vec{V}_{w_1}}{|\vec{U}_{w_1}| |\vec{V}_{w_1}|} \quad (3)$$

We also computed the similarity using Point-wise Mutual Information (PMI) for a must-link generation. PMI value is computed as shown in Equation (4).

$$\text{PMI}(w_1, w_2) = \log \frac{P(w_1, w_2)}{P(w_1)P(w_2)} \quad (4)$$

Finally, we combined the concise similarity with the PMI for checking the word relatedness. We computed the coherence between w_1 and w_2 using the following Equation (5).

$$\text{Coherence}(w_1, w_2) = \text{sim}(w_1, w_2) \text{PMI}(w_1, w_2) \quad (5)$$

Topic Modeling

We used the topic modeling Gibbs sampler algorithm to extract the topic based on automatically incorporating the domain knowledge enriched by global and local contexts. The overall process flow is depicted in Figure 8.

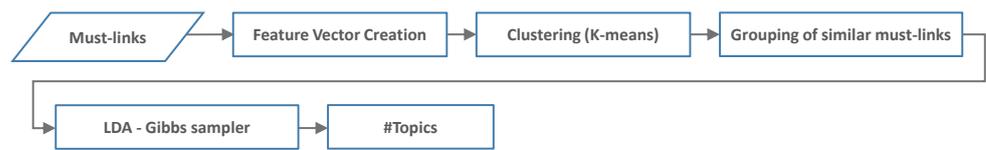


Figure 8. Work flow of integrating must-link into the Gibbs sampler.

UX Dimensions Generation

This section explains the process of UX dimension generation by auto labeling each extracted topic in the preceding section. We used the dictionary-based approach for classifying each topic based on top “n” words. The overall flow is depicted in Figure 9.

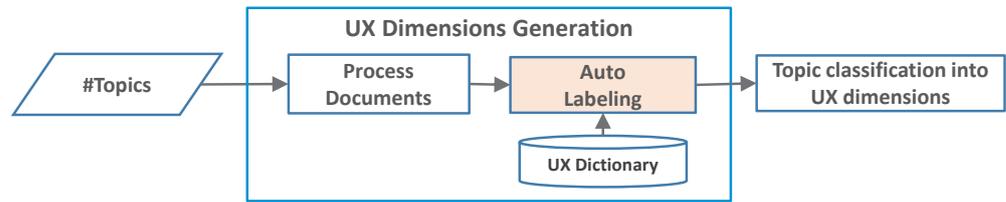


Figure 9. Workflow of topic labeling into UX dimensions.

We build the lexicon dictionary based on terms already used in previously validated scales [23,41,42] for measuring different aspects of UX using systematic review process. We selected the 223 terms, and then applied the WordNet for word expression. The final thesaurus contains 500 terms by adding the synonyms to the UX dictionary. Finally, we validated the UX dictionary using Cohen’s kappa coefficient [43] from three domain experts.

For topics classification based on the dictionary approach, we used the MeaningCloud text mining API. MeaningCloud allows developers to define the custom dictionary in the form of ontology. We created the UX dimensions dictionary with selected terms for topic labeling.

3.3. Sentiment Analyzer

For sentiment analysis, we employed the huggingface transformer sentiment analysis pipeline [44] in this work, which takes the set of online reviews ($R_i = \{r_1, r_2, \dots, r_n\}$) related to extracted UXDs. Based on the sentiment alignment of each review in R_i with extracted UXDs (the i th UXD), we generated the structured data, as shown in Table 2. We used the following equation x for sentiment orientation of D_i in online reviews R_i .

$$S_{mi}^* = \begin{cases} 1 & \text{if the sentiment orientation is } * \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

* represent the sentiment orientation, where $* \in \{pos, neg\}$ as shown in Table 2. The sentiment values are encoded to nominal as, if the sentiment of a review is positively associated with dimension, then $s_m^{pos} = 1$ and $s_m^{neg} = 0$; if the sentiment in the review is negative, then $s_m^{pos} = 0$ and $s_m^{neg} = 1$; if the sentiment in the review is neutral, then $s_m^{pos} = 0$ and $s_m^{neg} = 0$.

3.4. User Satisfaction Modeling

We previously discussed measuring the effect of user positive or negative sentiments toward each UXD on user satisfaction using a bidirectional LSTM model. The bidirectional LSTM model overcame the problems of the existing model used for user satisfaction models such as Gaussian distribution [13] and regression analysis.

Most of the existing models for user satisfaction assume that the online rating given by a user is a linear amalgamation of the sentiment regarding all the dimensions discussed

in the online reviews. However, this assumption is not valid; there are many issues, such as a complex combination of sentiments towards most of the dimensions in user online reviews. In order to resolve this issue, bidirectional LSTM outperforms when compared to other models for modeling user satisfaction. Based on this reason, we employed this model for measuring the sentiment effects towards each UXD. We used the user rating as label attributes for each review, along with generated data as discussed the in subsequent section for model training.

Table 2. Sentiment orientation toward each dimension.

Online Reviews	UX Dimensions (UXDs)							
	D_1		D_1		...	D_n		
	Pos	Neg	Pos	Neg		Pos	Neg	
r_1	1	0	0	1	...	0	0	
r_2					...	0	1	
...	
r_n	0	1	1	0	...	1	0	

Kano Model

We employed the Kano model, developed by Kano et al. [31], which is a two-dimensional model. Kano model is a well-known model of user satisfaction. This model categorizes product features into different classes including must-be, performance, excitement, indifferent, and reverse. These features values are associated with user satisfaction [14]. Details of each feature are described as follows:

1. Must-be: These features are essential customer’ requirements and expectations and are taken for granted. These features must be fulfilled, otherwise the product customer becomes dissatisfied.
2. One-dimensional (Performance): These features are related to product quality promised by the product service provider. These features have a direct impact on customer satisfaction when fulfilled.
3. Attractive (Excitement): These features give satisfaction when filled, but have no effect on customer dissatisfaction.
4. Indifferent: These product features neither influence on user satisfaction nor dissatisfaction.
5. Reverse: These features reveal a more significant degree of achievement, causing more customer dissatisfaction.

Based on the rules defined by [14], we also mapped the UXDs on the Kano model. We mapped the \vec{w}_i^{pos} and \vec{w}_i^{neg} on the Kano model’s five categories for modeling the user satisfaction. The details of these rules are the following:

1. If $\vec{w}_i^{pos} \leq 0$ and $\vec{w}_i^{neg} < 0$ then UXD_i is a must-be.
2. If $\vec{w}_i^{pos} \leq 0$ and $\vec{w}_i^{neg} \geq 0$ then UXD_i is a reverse.
3. If $\vec{w}_i^{pos} > 0$ and $\vec{w}_i^{neg} < 0$ then UXD_i is a performance.
4. If $\vec{w}_i^{pos} > 0$ and $\vec{w}_i^{neg} \geq 0$ then UXD_i is an excitement.

Figure 10 depicts the mapping on the Kano model based on the rules above.

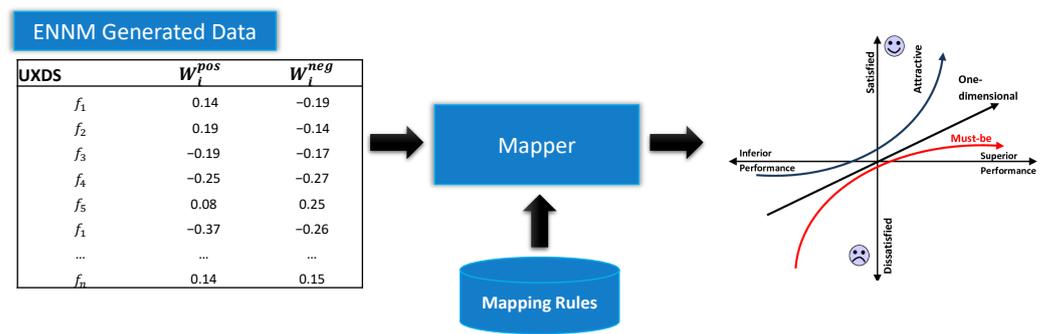


Figure 10. Mapping of UXDs on Kano Model.

4. Results and Case Study

To evaluate the efficiency of the proposed solutions, different experiments were performed at different levels. We used the dataset from [45], which contains the review data of both electronics and non-electronics products. Each domain category consists of 50 different products with a total of 1000 reviews. We evaluated the different parts of the proposed solution, such as (i) the experimental results and evaluations of part-1 (Data Collection) and (ii) the experimental results and evaluations of part-2 (User Review Analysis). A detailed explanation and the results are discussed in the following section.

4.1. User Review Quality Filters

We evaluated the user review quality filters models in terms of accuracy. We achieved 98.49% accuracy for training and 89.38% for model evaluation for “Spam Detection Classifier”. The overall performance of the “Relatedness Classifier” model was 95% for training and 90% for testing.

4.2. Topic Extractor

For topic modeling evaluation, we used the UMass topic coherence [28] metrics. The topic coherence (TC) metrics calculate the words’ relatedness within the topics; higher coherence values means a good topic. TC is computed as:

$$C(t; V^{(t)}) = \sum_{m=2}^M \sum_{l=1}^{m-1} \log \frac{D(V_m^{(t)}, V_l^{(t)})+1}{D(V_l^{(t)})} \tag{7}$$

In this section, we show an example of topics generated by UXWE-LDA, WE-LDA, and LDA to show an improvement by our proposed topic extractor. The red color in each topic in Table 3 show errors, as UXWE-LDA extracted more coherent and meaningful topics compared to the other baseline models.

We performed parameter tuning of UXWE-LDA such as number of top seed words (n), words similarity (m), and trust score (u). We examined the sensitivity of the three parameters of UXWE-LDA such as top seed words n, most similar words m, and the trust score u. The number of top 15 words gives us more coherent topic with higher TC value with other parameters setting for the given dataset. The experimental results reveals that a top 15 seed word gives higher TC value for electronic data, and for the non-electronic dataset, the top 25 seed words give us higher TC value. Therefore, we can conclude that a few seeds generate the coherent topic.

Additionally, from experimental results we reveal that TC values increase with more similar words at the initial stage, and gives us a higher TC value at 15 in the electronic dataset; for the non-electronic dataset, the model gives a higher value at 15.

Table 3. Example topics generated by UXWE-LDA, WE-LDA, and LDA. The bold text shows an error.

Electronics Products Dataset (TV-Screen)			Non-Electronics Products Dataset (Food)		
UXWE-LDA	WE-LDA	LDA	UXWE-LDA	WE-LDA	LDA
screen	image	screen	vegetable	flavor	popcorn
headset	video	side	cook	sweet	functions
microphone	end	big	meat	salt	bag
speaker	microsoft	line	pizza	market	potato
voice	logitech	flat	soup	fruit	healthy
sound	resolution	image	baked	natural	chicken
mic	top	top	delicious	spice	meat
audio	cd	lcd	chicken	strong	sweet
conversation	pc	resolution	bean	ingredient	number
mike	skype	sound	yummy	delicious	machine
loud	chat	strap	winter	bean	basic
bottom	picture	substitute	simple	open	spice

We examined the electronic products dataset and found TC increases with more similar words at the beginning, then plateaus and gives higher value at $m = 15$. For the non-electronic product dataset, the UXWE-LDA model gives an almost identical TC value and higher at $m = 25$. This shows that high quality knowledge is generated by must-links which are produced by the best seed words and word similarity. The similarity computation using TC ensures the quality of a must-link and that proper knowledge is incorporated into UXWE-LDA.

Figure 11 shows the average TC of each model using different number of topics on two datasets.

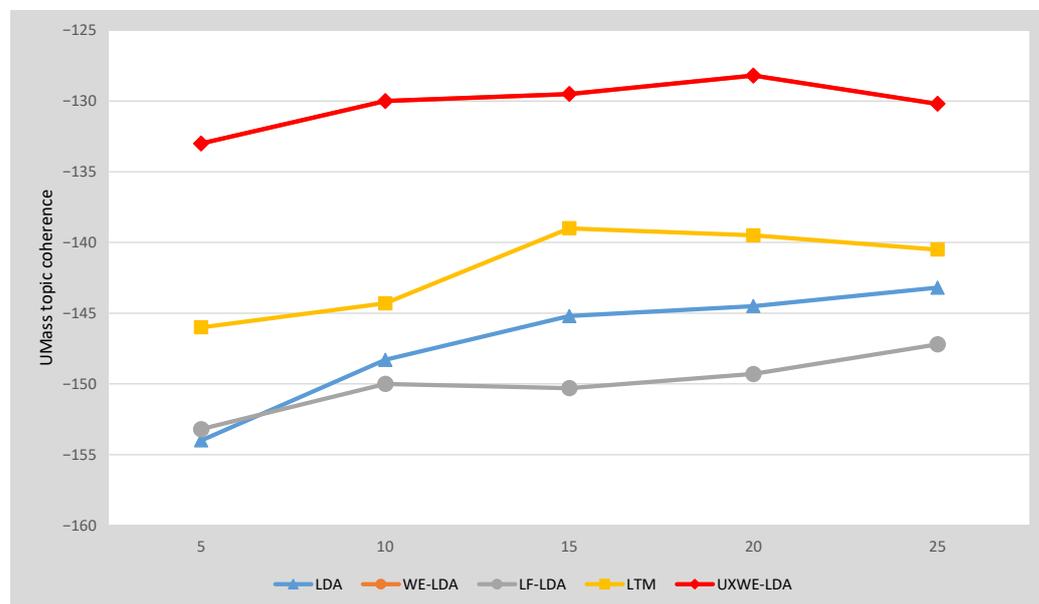


Figure 11. Average TC of top words with different number of topics on on both datasets.

The results show that, with the different number of topics and setting, UXWE-LDA always gives higher TC value than the other models, which shows that the UXWE-LDA is vigorous with a different combination of must-link clusters. Enhancements of UXWE-LDA over other models with p -value ($p < 0.007$) were significant in the two-tailed paired t -test.

4.3. Overall Comparison—Extrinsic UXDs Extraction Evaluation

We chose UX experts from a group working with us on an ongoing research project. We performed an extrinsic evaluation by comparing the UXWE-LDA inferred topic with the gold-label topic assigned by the three UX experts. The UX experts annotated a total of

300 online reviews, where each sentence is label based on the provided UX dimension list. Sentences with mutual agreement from all three annotators were considered as gold-label for the performance evaluation. We employed the topic-wise performance metrics (recall, precision, and F1 score) for comparison with LDA baseline algorithms. Precision means the percentage of correct classifications of that topic among all gold-label reviews sets, where the UXWE-LDA model predicts that topic. Recall for a topic is the portion of correct classifications of that topic out of all the cases of that topic in the gold-label reviews. The F1 score of a topic is the harmonic mean of recall and precision of that topic and is given in Equation (8).

$$\text{F1 Score} = 2 \times \frac{\text{precision}_k \times \text{recall}_k}{\text{precision}_k + \text{recall}_k} \quad (8)$$

where a higher F1 score indicates that the model performs well for classifying the test data as shown in Table 4.

Table 4. Topic-wise performance measures.

Topics	LDA			UXWE-LDA		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score
Attractiveness	0.71	0.46	0.55	0.83	0.72	0.77
Dependability	0.78	0.49	0.60	0.80	0.91	0.85
Efficiency	0.73	0.60	0.66	0.76	0.77	0.76
Perspicuity	0.80	0.47	0.59	0.80	0.72	0.76
Novelty	0.76	0.51	0.61	0.81	0.76	0.78
Stimulation	0.75	0.47	0.58	0.87	0.81	0.84

Figure 12 has shown that UXWE-LDA achieves 3% improvement on average as compared with LDA due to the incorporation of semantic domain knowledge.

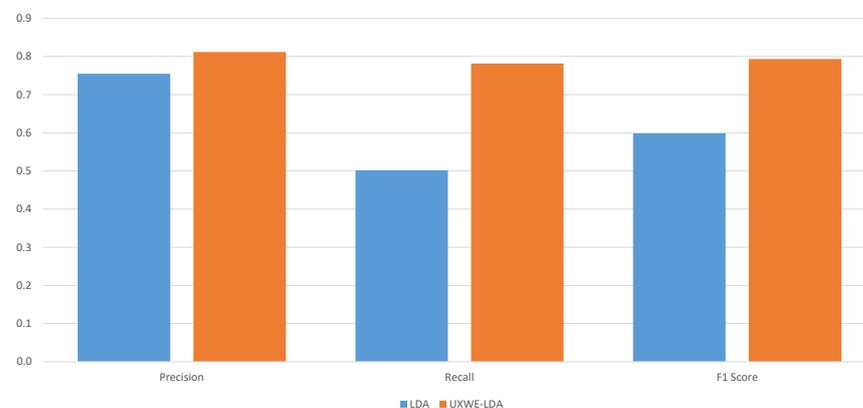


Figure 12. Average F1 Score, Precision, and Recall of LDA and UXWE-LDA.

4.4. UX Expert Base Evaluation

We compared the extracted dimensions using UXWE-LDA analysis with those manually extracted by our human experts for validation. We used the Jaccard coefficient similarity [46] to check the degree of dimensions overlapping between automatic extraction using UXWE-LDA and human experts. The Jaccard coefficient is calculated as Equation (9)

$$JC = \frac{|D_{UXWE-LDA} \cap D_{Exp}|}{|D_{UXWE-LDA} \cup D_{Exp}|} \quad (9)$$

where $D_{UXWE-LDA}$ dimensions are extracted using automatic UXWE-LDA analysis and D_{EXP} are dimensions extracted by human experts through a rigorous manual process. The higher the Jaccard coefficient's value, the higher the degree of overlap between the two sets

of dimensions, as shown in Table 5. Three researchers were invited, having hands-on NLP and text mining experience, to extract the UXDs from the randomly selected online reviews. Each researcher selected 50 reviews randomly; finally, a total of 150 reviews were selected for UXWE-LDA validation. We compared the UXDs extracted from UXWE-LDA with the UXDs extracted by the human experts to check the reliability of the result generated by UXWE-LDA.

The Jaccard coefficient for both the researchers’ extracts and UXDs extracted by UXWE-LDA model are 0.3, 0.5, and 0.4, respectively. This concludes that our study inferred new latent variables or dimensions from the online reviews. We claim that our study outcomes are more reliable for generalization due to large corpus textual data. Due to the complexity and ambiguity involved in UXD extraction from online reviews, the results show that UXWE-LDA is a reliable and suitable approach for UXD extraction from online reviews.

Table 5. A comparison of UXDs between UXWE-LDA model and human experts.

Dimensions	UXWE-LDA	Human Expert 1	Human Expert 2	Human Expert 3
Attractiveness	✓	X	✓	✓
Dependability	✓	✓	✓	✓
Efficiency	✓	X	✓	X
Perspiciuity	✓	X	✓	✓
Novelty	✓	✓	✓	✓
Stimulation	✓	✓	X	X
Aesthetics	X	✓	X	✓
Complexity	X	✓	✓	✓
Affect and emotion	X	X	✓	X

4.5. Case Study

We used publicly available Amazon data [47] of user reviews related to games reviews for this case study. The online reviews contain different words used by the different users to express their opinion; some words form “the long tail” as depicted in Figure 13. In total, 122,502 numbers of words were considered after applying the preprocess for UX dimensions extraction.

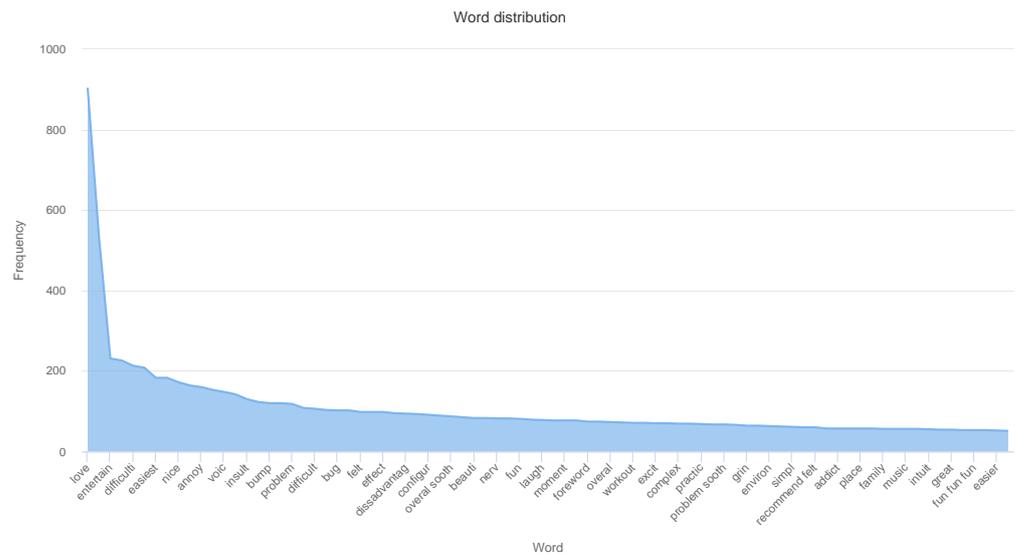


Figure 13. Word distribution for games reviews.

Figure 14 shows the frequency of user satisfaction score in terms of rating in the used dataset.

First, we applied the UXMCQ model to filter out the unrelated reviews. Then, we applied the UXWE-LDA model for the dimension extraction. Figure 15 shows the extracted dimensions from the online reviews.

The user sentiment orientations towards each UXD of online user reviews are shown in Figure 16. The results show that the users have positive opinions in the extracted UX dimensions as compared to negative.

We used the structured data with W_i^{pos} and W_i^{neg} vectors generated by part-2 of the proposed methodology to train the ENNM Model as shown in Table 6.

According to Table 6 generated by the ENNM model, the category of each UXD of game reviews can be identified and mapped in the Kano Model, as shown in Figure 8.

Figure 17 shows the threshold for determining whether a UXD is an indifferent UXD. As can be seen from Figure 17, the UXDs identified as excitement UXDs include: hedonic and perspicuity; pragmatic identified as reversed UXD; must-be UXD includes involvement; and efficiency. Finally, performance UXD consists of three dimensions (stimulation, attractiveness, and dependability).

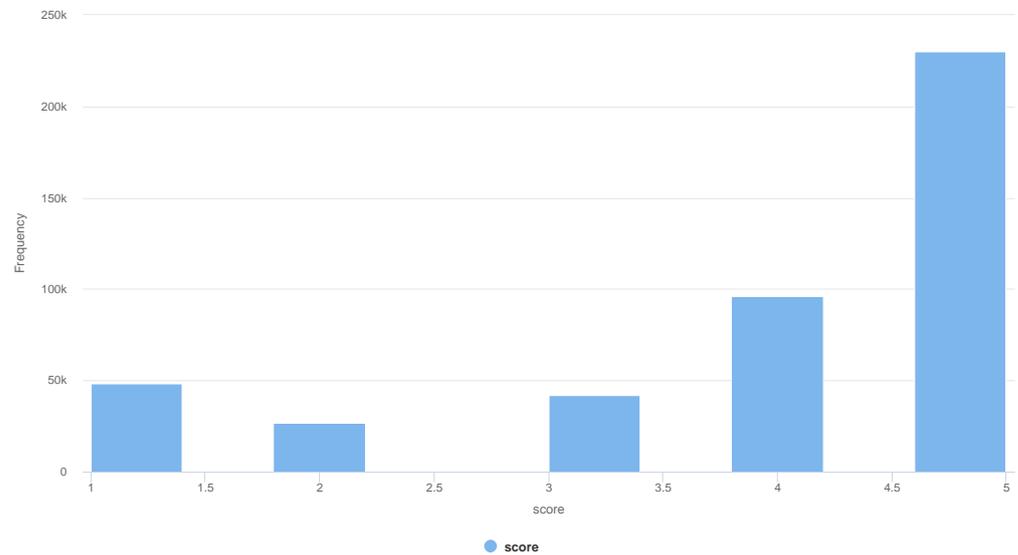


Figure 14. Overall user rating on reviews.

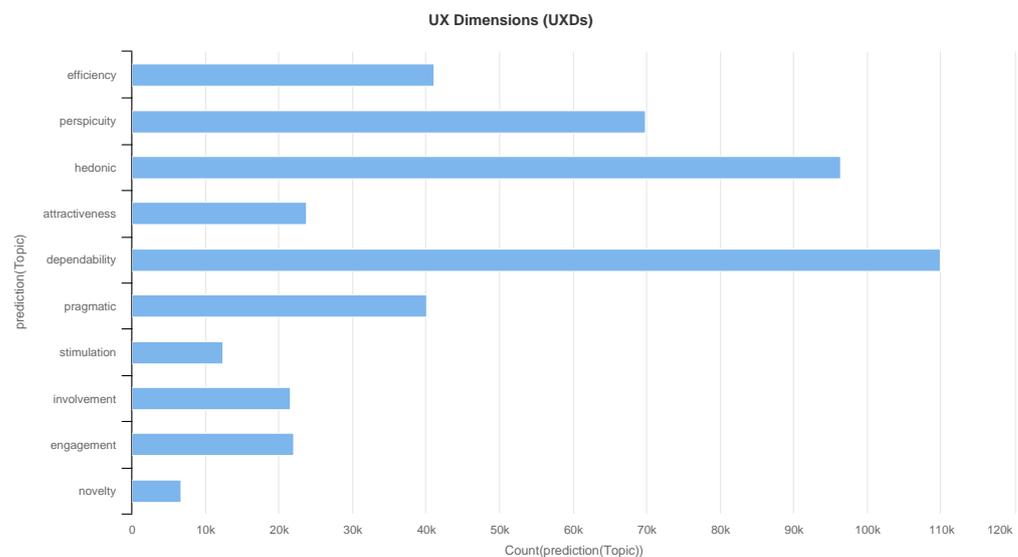


Figure 15. Extracted UX dimensions from user reviews.

Table 6. The values of positive and negative vectors generated by ENNM.

UXD	W_i^{pos}	W_i^{neg}
Attractiveness	0.14	-0.19
Dependability	0.19	-0.14
Efficiency	-0.19	-0.17
Engagement	-0.25	-0.27
Hedonic	0.08	0.25
Involvement	-0.37	-0.26
Perspicuity	0.14	0.15
Pragmatic	-0.18	0.04
Stimulation	0.03	-0.08

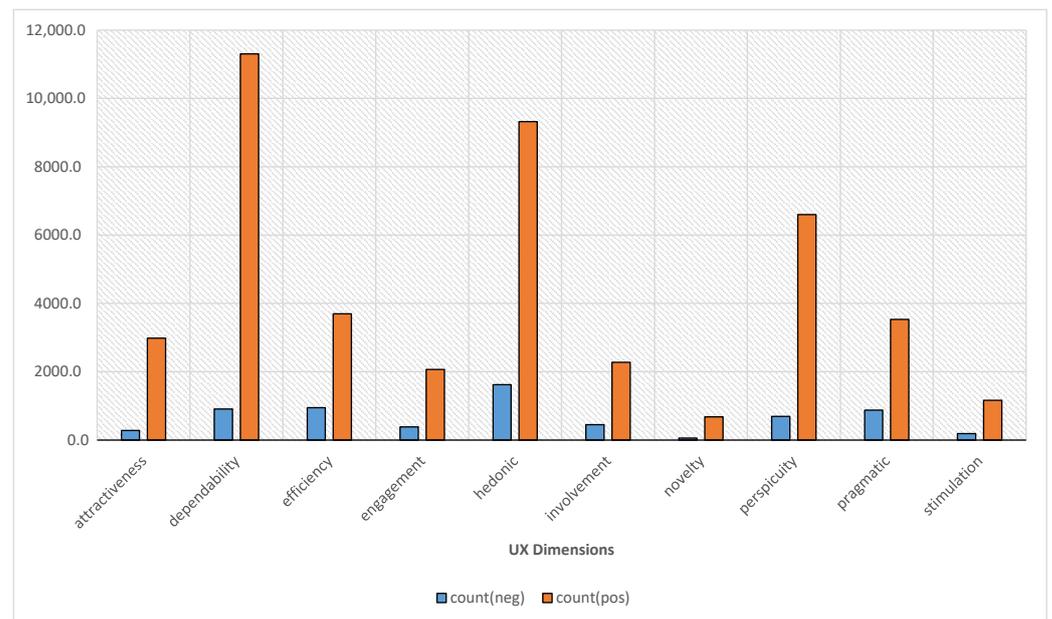


Figure 16. The sentiment orientations result towards each extracted UXD.

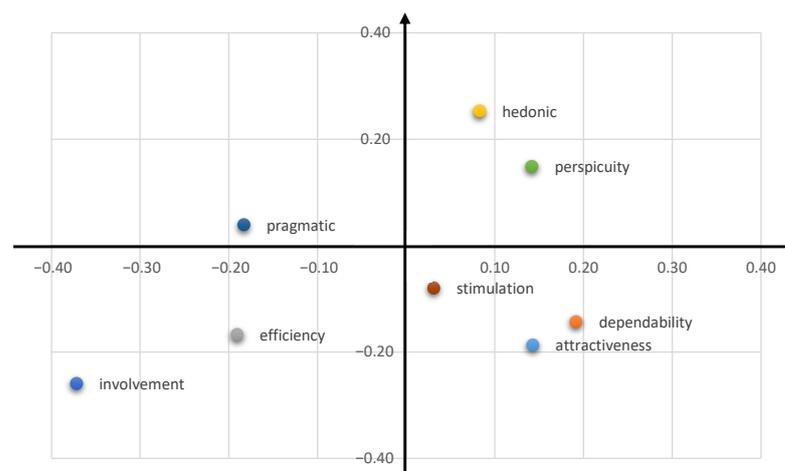


Figure 17. Mapping the extracted dimensions IN THE Kano Model.

5. Discussion

This study proposes a comprehensive data-driven strategy for incorporating UX factors into UX modeling using online reviews. The results of this study may assist product designers in overcoming obstacles faced by current UX studies, particularly in terms of how

UX research data are acquired, which are essential for UX research but time-consuming and may not be comprehensive. Studies show that sentiment analysis is substantially impacted by product features, the context of use, and user cognitive aspect, and a solution to the issue of entity detection and assignment for opinion mining applications HAS been provided. Based on these past investigations, we have investigated a systematic, data-driven method for incorporating UX dimensions into UX modeling from online user reviews.

Our study has expanded earlier UX models with additional UX characteristics (such as hedonic and pragmatic qualities) and explains how this might be implemented. In this study, we use NLP approaches to increase the list of terms describing UX aspects before automatically extracting UX aspect data from online user reviews. We present a case study to show how UX-relevant data can be collected from online user reviews. The results are consistent with what domain experts have recommended. This case study shows that our methodology can find UX aspect data and enable UX analysis. In the meantime, we also observe that several areas need additional investigation.

To assess whether a statement is connected to a specific UX component (hedonic or pragmatic), we first used a variety of filter-based algorithms on online user reviews.

Second, we employed an unsupervised approach to mining UXDs using state-of-the-art techniques by incorporating the UX domain knowledge. This research examined the strategic and tactical actions of proactive thinking and UX design idea development.

The presented study has potential implications in product design, such as it can be used to mine the user opinion toward each UX aspect so that product designers can make a better decision to improve the positive UX of their customers. Additionally, they can further know the strengths and weaknesses of the product. This method also allows the product designer to understand the different categories of UXDs in terms of the Kano model, which is essential for product enhancement. According to the classification results of UXDs, the priority order of UXDs for developing product enhancement plans can be determined. Designers may utilize the data to analyze a product's position above competitors regarding features and UX, allowing potential product modifications. UX study findings enable companies to explore new markets and enhance commercial decision making, from segmenting prospective consumers to strategic product design.

6. Conclusions

Due to advancements in social media platforms, users post their opinions in the form of online reviews daily. These online reviews contain beneficial information related to UX. The online user reviews can be used to understand UX and UX modeling. This study developed a data-driven methodology to mine UX-related information from these substantial online reviews. The automatic approach overcomes the problems of manually analyzing those vast data. We designed and verified a machine-learning-based computational method for mining the UXDs for UX modeling from online user reviews.

In the method, first, we filter those reviews unrelated to the UX domain using UX multi-criteria qualifiers (UXMCQ). Then, we extract the UXDs from the filtered reviews using an enhanced topic extraction methodology called UXWE-LDA. UXWE-LDA improves the existing knowledge-based topic models by extracting more domain-dependent dimensions in the UX area through UGC. Finally, user satisfaction was modeled using the Kano model by mapping the UX dimensions.

However, our results would be more accurate if we could integrate reviews from other sources' databases, and we will address this in our future study to enhance UX assessment for products and services. Furthermore, the study neglected some words, such as infrequent words, which help indicate user preference and needs for a product or service. Therefore, we need to examine effective solutions for incorporating word embedding. Furthermore, the experiment also needs to be extended with other settings and datasets to overcome the existing limitations.

Author Contributions: Conceptualization, J.H.; methodology, J.H., software, J.H.; validation, J.H., M.A. and S.L.; investigation, J.H. and M.R.; resources, J.H., M.A. and S.L.; data curation, M.R. and

H.F.A.; writing—original draft preparation, J.H., Z.A. and H.F.A.; writing—review and editing, J.H., Z.A., M.A. and H.F.A.; visualization, J.H.; supervision, S.L.; project administration, S.L.; and funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-0-01629) supervised by the IITP (Institute for Information & communications Technology Promotion), by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-00655), by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2020-0-01489), (IITP-2021-0-00979) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2022-0-00078, Explainable Logical Reasoning for Medical Knowledge Generation).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, C.E.; Lai, Y.H. Quasi-ADS-B Based UAV Conflict Detection and Resolution to Manned Aircraft. *JECE* **2015**, *2015*, 297859. [[CrossRef](#)]
2. Pucillo, F.; Cascini, G. A framework for user experience, needs and affordances. *Des. Stud.* **2014**, *35*, 160–179. [[CrossRef](#)]
3. Law, E.L.C.; Roto, V.; Hassenzahl, M.; Vermeeren, A.P.; Kort, J. Understanding, Scoping and Defining User Experience: A Survey Approach. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; Association for Computing Machinery: New York, NY, USA, 2009; pp. 719–728. [[CrossRef](#)]
4. Hassenzahl, M. User Experience (UX): Towards an Experiential Perspective on Product Quality. In Proceedings of the 20th Conference on l’Interaction Homme-Machine, Metz, France, 2–5 September 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 11–15. [[CrossRef](#)]
5. *ISO 9241-11: 2018; Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts.* ISO: Geneva, Switzerland, 2018.
6. Law, E.L.C.; van Schaik, P. Modelling user experience—An agenda for research and practice. *Interact. Comput.* **2010**, *22*, 313–322. [[CrossRef](#)]
7. Hussain, J.; Khan, W.A.; Hur, T.; Bilal, H.S.M.; Bang, J.; Hassan, A.U.; Afzal, M.; Lee, S. A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation. *Sensors* **2018**, *18*, 1622. [[CrossRef](#)] [[PubMed](#)]
8. Kosmadoudi, Z.; Lim, T.; Ritchie, J.; Louchart, S.; Liu, Y.; Sung, R. Engineering design using game-enhanced CAD: The potential to augment the user experience with game elements. *Comput.-Aided Des.* **2013**, *45*, 777–795. [[CrossRef](#)]
9. Law, E.L.C.; van Schaik, P.; Roto, V. Attitudes towards user experience (UX) measurement. *Int. J. Hum.-Comput. Stud.* **2014**, *72*, 526–541. [[CrossRef](#)]
10. Khan, J.; Alam, A.; Hussain, J.; Lee, Y.K. EnSWF: Effective features extraction and selection in conjunction with ensemble learning methods for document sentiment classification. *Appl. Intell.* **2019**, *49*, 3123–3145. [[CrossRef](#)]
11. Yang, B.; Liu, Y.; Liang, Y.; Tang, M. Exploiting user experience from online customer reviews for product design. *Int. J. Inf. Manag.* **2019**, *46*, 173–186. [[CrossRef](#)]
12. Guo, Y.; Barnes, S.J.; Jia, Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tour. Manag.* **2017**, *59*, 467–483. [[CrossRef](#)]
13. Farhadloo, M.; Patterson, R.A.; Rolland, E. Modeling customer satisfaction from unstructured data using a Bayesian approach. *Decis. Support Syst.* **2016**, *90*, 1–11. [[CrossRef](#)]
14. Bi, J.W.; Liu, Y.; Fan, Z.P.; Cambria, E. Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model. *Int. J. Prod. Res.* **2019**, *57*, 7068–7088. [[CrossRef](#)]
15. Vu, H.Q.; Li, G.; Law, R.; Zhang, Y. Exploring Tourist Dining Preferences Based on Restaurant Reviews. *J. Travel Res.* **2019**, *58*, 149–167. [[CrossRef](#)]
16. Suryadi, D.; Kim, H.M. A Data-Driven Approach to Product Usage Context Identification From Online Customer Reviews. *J. Mech. Des.* **2019**, *141*, [[CrossRef](#)]
17. Folmer, E.; van Gurp, J.; Bosch, J. A framework for capturing the relationship between usability and software architecture. *Softw. Process. Improv. Pract.* **2003**, *8*, 67–87. [[CrossRef](#)]
18. Seffah, A.; Donyaee, M.; Kline, R.B.; Padda, H.K. Usability measurement and metrics: A consolidated model. *Softw. Qual. J.* **2006**, *14*, 159–178. [[CrossRef](#)]

19. Hornbæk, K. Current practice in measuring usability: Challenges to usability studies and research. *Int. J. Hum.-Comput. Stud.* **2006**, *64*, 79–102. [CrossRef]
20. Leem, B.H.; Eum, S.W. Using text mining to measure mobile banking service quality. *Ind. Manag. Data Syst.* **2021**, *121*, 993–1007. [CrossRef]
21. von Wilamowitz-Moellendorff, M.; Hassenzahl, M.; Platz, A. Dynamics of user experience: How the perceived quality of mobile phones changes over time. In Proceedings of the User Experience-Towards a Unified View, Workshop at the 4th Nordic Conference on Human-Computer Interaction, Oslo, Norway, 14–18 October 2006; pp. 74–78.
22. Ferreira, D.J.; Melo, T.F.; do Carmo Nogueira, T. Unveiling Usability and UX Relationships for Different Gender, Users Habits and Contexts of Use. *J. Web Eng.* **2020**, *19*, 819–848. [CrossRef]
23. Hassenzahl, M.; Burmester, M.; Koller, F. AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. In *Mensch & Computer*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 187–196.
24. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *Int. J. Interact. Multimed. Artif. Intell.* **2017**, *4*, 40–44. [CrossRef]
25. Tirunillai, S.; Tellis, G.J. Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *J. Mark. Res.* **2014**, *51*, 463–479. [CrossRef]
26. Xu, F.; La, L.; Zhen, F.; Lobsang, T.; Huang, C. A data-driven approach to guest experiences and satisfaction in sharing. *J. Travel Tour. Mark.* **2019**, *36*, 484–496. [CrossRef]
27. Chang, J.; Gerrish, S.; Wang, C.; Boyd-graber, J.; Blei, D. Reading Tea Leaves: How Humans Interpret Topic Models. In *Advances in Neural Information Processing Systems*; Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C., Culotta, A., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2009; Volume 22.
28. Mimno, D.; Wallach, H.M.; Talley, E.; Leenders, M.; McCallum, A. Optimizing Semantic Coherence in Topic Models. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, 27–31 July 2011; Association for Computational Linguistics: Stroudsburg, PA, USA, 2011; pp. 262–272.
29. Yao, L.; Zhang, Y.; Chen, Q.; Qian, H.; Wei, B.; Hu, Z. Mining coherent topics in documents using word embeddings and large-scale text data. *Eng. Appl. Artif. Intell.* **2017**, *64*, 432–439. [CrossRef]
30. Decker, R.; Trusov, M. Estimating aggregate consumer preferences from online product reviews. *Int. J. Res. Mark.* **2010**, *27*, 293–307. [CrossRef]
31. Matzler, K.; Hinterhuber, H.H.; Bailom, F.; Sauerwein, E. How to delight your customers. *J. Prod. Brand Manag.* **1996**, *5*, 6–18. [CrossRef]
32. Ott, M.; Choi, Y.; Cardie, C.; Hancock, J.T. Finding Deceptive Opinion Spam by Any Stretch of the Imagination. *arXiv* **2011**, arXiv:1107.4557. <https://doi.org/10.48550/arXiv.1107.4557>.
33. Maiya, A.S. ktrain: A Low-Code Library for Augmented Machine Learning. *J. Mach. Learn. Res.* **2020**, *23*, 1–6. [CrossRef]
34. Hussain, J.; Lee, S. *Review-Based User Experience (UX) Modeling*; The Korean Institute of Information Scientists and Engineers: Seoul, Korea, 2016; pp. 980–982.
35. Loria, S. Textblob Documentation. Release 0.15. 2018. Volume 2, p. 269. Available online: <https://textblob.readthedocs.io/en/dev> (accessed on 10 March 2022)
36. Hamilton, W.L.; Clark, K.; Leskovec, J.; Jurafsky, D. Inducing domain-specific sentiment lexicons from unlabeled corpora. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–5 November 2016; Volume 2016, p. 595.
37. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global Vectors for Word Representation. In Proceedings of the Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
38. Shi, T.; Kang, K.; Choo, J.; Reddy, C.K. Short-Text Topic Modeling via Non-Negative Matrix Factorization Enriched with Local Word-Context Correlations. In Proceedings of the 2018 World Wide Web Conference, Lyon France, 23–27 April 2018; pp. 1105–1114. [CrossRef]
39. Viegas, F.; Canuto, S.; Gomes, C.; Luiz, W.; Rosa, T.; Ribas, S.; Rocha, L.; Gonçalves, M.A. CluWords: Exploiting Semantic Word Clustering Representation for Enhanced Topic Modeling. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, Australia, 11–15 February 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 753–761. [CrossRef]
40. Rong, X. Word2vec Parameter Learning Explained. *arXiv* **2014**, arXiv:1411.2738.
41. Lavie, T.; Tractinsky, N. Assessing dimensions of perceived visual aesthetics of web sites. *Int. J. Hum.-Comput. Stud.* **2004**, *60*, 269–298. [CrossRef]
42. Laugwitz, B.; Held, T.; Schrepp, M. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*; Holzinger, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 63–76.
43. Kraemer, H.C. Kappa coefficient. In *Wiley StatsRef: Statistics Reference Online*; Wiley: Hoboken, NJ, USA, 2015; pp. 1–4. [CrossRef]
44. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-Art Natural Language Processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, 16–20 November 2020; pp. 38–45. [CrossRef]

45. Chen, Z.; Liu, B. Mining Topics in Documents: Standing on the Shoulders of Big Data. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; Association for Computing Machinery: New York, NY, USA, 2014; pp. 1116–1125. [\[CrossRef\]](#)
46. Niwattanakul, S.; Singthongchai, J.; Naenudorn, E.; Wanapu, S. Using of Jaccard coefficient for keywords similarity. In Proceedings of the International Multiconference of Engineers and Computer Scientists, Hong Kong, 13–15 March 2013; Volume 1, pp. 380–384.
47. McAuley, J.; Leskovec, J. Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text. In Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, 12–16 October 2013; Association for Computing Machinery: New York, NY, USA, 2013; pp. 165–172. [\[CrossRef\]](#)