Decision Support Systems 162 (2022) 113864

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Mining longitudinal user sessions with deep learning to extend the boundary of consumer priming

Li-Chen Cheng [a,*], Kuanchin Chen [b,c]

[a] Department of Information and Finance Management, National Taipei University of Technology, Taipei 106, Taiwan, ROC
[b] Department of Business Information Systems, Western Michigan University, Kalamazoo, MI, USA
[c] Center for Business Analytics, Western Michigan University, Kalamazoo, MI, USA

ARTICLE INFO

Keywords:
Text mining
Session-based recommendation
Priming theory
E-commerce

ABSTRACT

Priming is challenging when consumers start shortlisting products before the final purchase. This is because this shortlisting process is performed in multiple user sessions online across time, the shortlist does not stay as a static list, and product comparison in this stage uses the heuristics internal to individual consumers. The goal of this study is two folds: (1) to approximate user heuristics after B&B product shortlisting using NLP and deep learning techniques, and (2) to identify optimized deep learning models for the representation of key elements of consumer heuristics. This offers an extension of the priming theory into product comparison and shortlisting stages that were traditionally difficult for marketers to tap into. By analyzing the B&B product information repeated visited in user sessions, the formation of shortlists is identified and products in the shortlists can then be compared. Subsequent priming and promotions can therefore be performed closer to the actual purchase. Our study also provides marketers keywords and their associated activated words relevant for crafting marketing messages. As these activation words are extracted from the B&B sites and product reviews that the users had visited repeatedly in long-term tracking sessions, they are analogous to effects produced from user participatory design, an approach popular in the IT world. Our work shows that opportunities for marketing decision support, especially into the shortlisting phase, are now possible through machine learning techniques. Both theoretical and practical implications are provided.

1. Introduction

Message priming refers to an approach to make a message more accessible in order to influence responses of message receivers [1]. For example, the primed message “Only pay for what you need” by Liberty Mutual Insurance activates the semantic importance of price individualization later when a consumer purchases a car insurance. A longer form of message priming includes the primed message either as a standalone message or as an embedded message into product description, ads, and other consumer-facing materials. Priming encourages brand identity, retention [2], customer loyalty [3,4] and purchase [5]. Priming effects weaken as the time goes by [1,2], but stronger if presented closer to the purchase decisions [5], such as during product shortlisting before the final purchase.

Because shortlisted products tend to be similar in product attributes that were used to select products to include in the shortlist, consumers usually compare the shortlisted products using additional supplemental information (e.g., usage experience in product reviews, additional product features, etc.) before making the final purchase [6]. Unlike commodity products where shortlisted items selected into the shopping carts are known to the vendors and certain last-minute marketing strategies (e.g., dynamic pricing in shopping carts [7]) could be applied as a result, such opportunity to tap into consumer shortlists for hotel booking is virtually non-existent for reasons such as consumers may not login when they start shortlisting, shortlisted products may not reside on vendor's platform (such as in the form of shopping carts), consumers may not rely on vendor's platform to retain their shortlists, and hotel comparison is done through micro-increments across different time [8]. Coupled with the restriction of the HTTP protocol being stateless where web servers treat every user visit as an independent, different user visit even if it was made by the same user at different times. As a result, linking consumer visits at different times even when they had not logged in is important to the success of shortlist identification. Without being able to identify the formation of shortlists in hotel bookings, the effects

* Corresponding author.
E-mail addresses: lijen.cheng@gmail.com (L.-C. Cheng), kc.chen@wmich.edu (K. Chen).

https://doi.org/10.1016/j.dss.2022.113864
Received 29 July 2021; Received in revised form 13 July 2022; Accepted 16 August 2022
Available online 6 September 2022
0167-9236/

of marketing interventions (such as priming and other promotion strategies), identification of competition during shortlisting, last-minute promotion with the right primed message and primed keywords to promote product differentiation, are either limited, difficult or impossible. For example, shortlisted products are quite homogeneous in their product attributes. Without knowing who are among the competitors during consumer shortlisting, priming or promotions emphasizing product attributes similar to the shortlisted items will unlikely be effective. Therefore, using the right priming keywords that effectively differentiate one's product from the competition is crucial after product shortlisting.

The predominant approaches to study priming effects have been simulation and lab-controlled experiments. Despite useful in delineating construct relationships, opportunities for generalizability and scalability for individualization could be limited [9]. As Jones & Chen indicated, the final shortlist is not always a static list of selections [6]. It could expand or contract as a consumer gains more experience, has more information regarding the product or revises their heuristics. The static results produced from simulation and lab studies could be too rigid for practical use [3] and costly to become scalable longitudinally.

To accomplish shortlist identification as it expands and contract, identification of primed keywords, and distilling of consumers' long-term preferences from a series of micro-increments of hotel selection [8], one would need a framework that breaks the technical and managerial boundaries by modeling long-term user preference, identifying shortlist formation and items within, and extracting primed keywords for shortlisted items from their product reviews. To the best of our knowledge, this is an area that has profound marketing opportunities, but has received very little attention in scholarly works. Therefore, this study is designed to take an initial look at these issues by answering the following key research questions:

RQ1. How to build a framework that allows for the identification and monitoring of product shortlisting BnBs?

RQ2. How to capture long-term user preference with deep learning and NLP techniques to recommend BnBs?

RQ1 is a theoretical advancement by extending the boundary of the priming theory into a product shortlisting phase that is internal to individual consumers. We proposed a framework that uses natural language processing (NLP) to extract consumer shortlisting longitudinally throughout user sessions across time and builds deep learning models to capture consumer heuristics during the product selection and purchase decisions. The framework produces three interesting deliverables for practitioners: (a) A recommendation system built based on the learned consumer heuristics to recommend B&B options, (b) identification of shortlists, and (b) keywords relevant to user heuristics extracted from repeat visits for concurrent and subsequent priming. Because these keywords are from the B&B materials that consumers repeated visited, they create a stronger appeal than the primed keywords solely defined by marketers. Not only does this approach allow for subsequent priming to be performed after shortlisting, it also provides opportunities for automating heuristics extraction to tackle the issue of shortlist not being static over time [6]. To answer RQ2, we propose a novel multi-layer solution to optimize deep learning and NLP techniques to capture consumer preferences from the BnB pages they visited and dimensions of the product reviews associated with the visited BnBs. We also present technical challenges of using deep learning techniques in the context of BnB booking and possible solutions.

In section 2, we highlight the current literature for developing a two-stage product evaluation process, and the current development of NLP and recommendation systems. Section 3 outlines our approach to empirically validate the two-stage process and proposes empirical analyses. Section 4 reports findings from the empirical validation. Discussion and conclusion are presented in section 5 before our contributions are outlined.

2. Related work

2.1. Priming and its boundary

Priming theories [10–12] suggest that a primed message triggers a network of associated words internal to an individual. In a typical priming activity (a.k.a., a priori priming [5]), a prime (in the forms of text, color, etc.) or sometimes called a "stimulus" is presented to the intended audiences, and the resulting goal is measured at a later time. Priming has been used heavily in marketing where forms of advertisement (text, color, etc.) are used to activate a network of associated words or knowledge about the product in consumers to eventually drive a result (e.g., purchase, brand recognition, emotions, etc.) [10].

However, using traditional a priori priming in an online environment meet several challenges. First, a priori priming requires the subject to be available during the time a stimulus is applied and later when the subject's execution of a task is observed [13]. In reality, this approach does not scale well even if the consumers are willing to participate in the experiment [14]. Research begins to propose a concept called concurrent priming, where the stimulus and task performance are presented at the same time [13]. As Minas and Dennis [13] shows, consumers under concurrent priming can still be easily distracted when other stimuli from competitors are available. This is the case in online shopping where competitors present similar or even stronger primes that could distract a consumer away from the original prime [15]. Most importantly, there is no guarantee that consumers even actually see the original prime [5]. Therefore, solving the scalability problem in a priori priming in online shopping by replacing it with concurrent priming also creates problems of its own. Second, effects of priming decay as the time goes by, therefore, it is most effective when it is done closer to the beginning of the intended task (e.g., purchase) [13]. Therefore, the time when consumers are shortlisting products before making a purchase is ideal for priming. However, the start of consumer shortlisting is not always visible to markers, especially in those products (such as hotels, air travel, resort booking, tours, etc.) where shopping carts or similar measures are typically unavailable for vendors to identify shortlisted items. Third, shortlists expand and contract across different time [3]. Compounded with this is that consumers tend to employ a different set of attributes than the original ones used for initial screening in order to evaluate shortlisted products before making the final purchase [6]. Selection of primed keywords to differentiate one's offering from the competitors' is increasingly important at this shortlisting stage.

The lack of clarity in the shortlisting process (i.e., the start of shortlisting, competition in shortlisted products, etc.) has limited the effectiveness of priming. Not only that it is unclear about the timing for the vendors to determine when subsequent priming to differentiate products should be effectively performed during shortlisting, not knowing the competition in the shortlist further limits the available marketing strategies until it is too late. Therefore, these challenges mark the boundary of the priming theory for the hotel industry. This is the reason Minton, Cornwell & Kahle [10] suggested researching into the boundary conditions of priming theories after reviewing the theoretical development in the priming literature. In the present study, we propose a deep learning approach with NLP to identify shortlisting in BnB booking by observing repeat visits (see how repeat visits leads to final purchase and other intended goals in [16]) and extract key priming keywords from the BnBs that consumers actually visited. By modeling the actual visits through a long-term manner, we are able to better address the challenges of priming in the online environment.

2.2. NLP and deep learning as a potential to extend the boundary of priming for hospitality industry

The product review literature is full of studies using natural language processing (NLP) [17–19] to uncover patterns from product reviews [18] to improve products or services and hotel revisits [20–22] and fraudulent review detection [23,24].

Several challenges arise when using existing NLP approaches to tap into consumer heuristics to extend the boundary of priming into the product shortlisting stage. First, although mining all product reviews without discretion (e.g., [17,25]) provides nice generalized insights, not all product reviews are visible or of interest [26] to consumers. Nor are all reviews equally weighted by the consumers. For example, attribute-based information is more relevant to screen products before shortlisting, but experienced-based or other additional information is more useful after shortlisting [6,27]. Although non-discretional selection of product reviews is still the main stream approach in product review studies. (e.g., [28]), it may not model preferences of individual consumers well, thus reducing the accuracy of recommender systems and opportunities for customization of consumer experiences. Second, reviews that are visited repeatedly by consumers weigh more to compare similar products in the shortlist [16]. This offers opportunities to tap into consumer heuristics when processing shortlisted products. Third, click stream data is difficult to aggregate across time for different visits of the same person because HTTP is a stateless protocol. Not all users will login to provide their identity when viewing products and their reviews. Fourth, product aspects important to an individual consumer may not be easily identified or measured with aggregated or researcher-defined components. For example, unidimensional scores (such as product rating) are known to be inaccurate or inconsistent compared with the written reviews [29]. Decomposing it into finer dimensions, each capturing a unique aspect of the same concept, enables deeper insights. For example, decomposing the quality score from hotel ratings into individual dimensions of accuracy, cleanliness, check-in, communication, location and value as in [30] allows businesses to address the finder issues relating to the target concept (e.g., quality). However, researcher-defined decomposition may not be ready for use in a priming message or promotions to have an individualized effect due to differences across individuals and weights of these dimensions at different stages of project selection. For example, location might matter during the initial hotel screening, but other experience aspects matter more after shortlisting.

Each of these challenges point to the limitation of non-discretional selection of product reviews, but pairing long-term clickstream data to identify visited reviews and sites for NLP provides opportunities to model actual use and shortlist identification. To the best of our knowledge, these areas have received little attention in scholarly works, but addressing them offers opportunities for business decisions. Therefore, we propose a framework of NLP and deep learning on long-term user sessions to shortlists and user preferences from the actual user clickstream data.

2.3. Deep learning systems and recommender systems

Recurrent neural network (RNN) and its improved forms - long-short term Memory (LSTM) and gated recurrent unit (GRU), have recently emerged as deep learning techniques to predict sequentially ordered data with significantly improved prediction accuracy [31,32]. Examples include modeling user sessions with RNN and gated recurrent units, GRU4rec, can be used to predict the probability of subsequent item views. It has also been used in item-based- k-nearest-neighbor (kNN) [33], and session-based nearest neighbor [34].

Three forms of recommendation systems are popular: content-based methods, collaborative-based filtering systems and hybrid systems [35,36].Content-based filtering methods recommend information (e.g., products) based its similarity to an individual's past preference [32]. Collaborative-based systems makes recommendations based on the preferences of similar users [36], but they suffer from rating sparsity, scalability, and efficiency issues [36]. The assumption of these classic recommendation systems is that they treat historical user activities equally compared to more recent activities. Recently, hybrid recommendation systems have emerged [33,36,37]. For example, session-based recommender systems (SBRSs) emerged as a hybrid approach to consider more on recent user activities. This is especially useful since more recent activities are more likely to be relevant to the consumer's recent preference [31,32]. Early SBRS approaches rely high frequency items to predict the user's next action [32]. Despite the importance, high frequency items may not be relevant to the consumer's task at hand. More recent development of SBRS include various methods based on Markov models [13]. One downside of Markov chain based recommendation systems is that it could only consider user's short sequence requiring data to be available to make the prediction for the next possible steps [31–33]. For the above reasons, both pattern mining based models and Markov models may not be the best types of recommender systems to mine and predict long-term discrete sequences (e.g., website visits). In recent years, deep learning is used together with SBRSs to fill the above gap with a better prediction accuracy [31,33]. The advantage of deep learning models is its powerful capabilities of handling different forms and time-based longitudinal session data to learn users' long- and short-term preferences. Traditional GRU based systems and RNN make a one-way pass through sessions treating all sessions equally [38,39]. Wang et al.'s survey of recommender systems attention-based systems [32] and Li, et al. [40] show that the incorporation of attention mechanism to focus on certain parts of the input sequence (e.g., information repeatedly visited) and algorithms that consider both long and short-term sequences (e.g., through LSTMs) are one key to improve the quality of the recommender systems. We follow this development and incorporate bidirectional LSTM that runs through the sequences in both directions to further enhance the quality proposed in these existing studies.

3. The system framework

AsiaYo is one of the largest online accommodation booking platforms in Asia.[1] They have serviced more than 600,000 customers since March 2014 across 60,000 B&Bs in more than 60 Asian cities. There are on average 2.5 million unique visits to their B&B platform each month. After reviewing user browsing history with the company, we see empirical patterns of two-stage B&B evaluation: screening to match the consumer's baseline requirements, and deep product comparison by reading customer reviews. This two-stage process is illustrated in part A of Fig. 1 below.

Part A of Fig. 1 illustrates the customer's two-stage product evaluation process. Numbered orange bubbles denote the key points in time to perform priming. Plenty of studies already reported priming to engage consumers at points one and four. Point one is initial product screening, but this could span across multiple user sessions, which makes tracking the same user challenging due to the stateless nature of the HTTP protocol. This is especially true when browsing B&Bs is usually anonymous without needing a user to login. Traditional priming stimuli and verification of them through task stimuli later on are difficult in point two and three because they are internal to the consumers, and most consumers do not have the time or willingness to spend time to verify the primed effects with the marketers [5]. What works in a lab setting can no longer function in an online environment. Label four is post-purchase services, and promotions.

Without being able to tap into steps two and three in Part A of Fig. 1 above is a loss of opportunity for marketers to perform further priming or promotion. Therefore, we turned to technical session aggregation techniques provided through AsiaYo that identify consumers by incorporating information regarding devices used for the visit, user identifiable information, internally generated identification, 3rd party cookies/sessions, and IP, into a holistic user identification strategy. We are only permitted to describe the high-level idea without revealing their trade secret, but this user identification approach is powerful enough to link

[1] We signed a non-disclosure agreement (NDA) with AsiaYo to analyze their consumer data for this research.

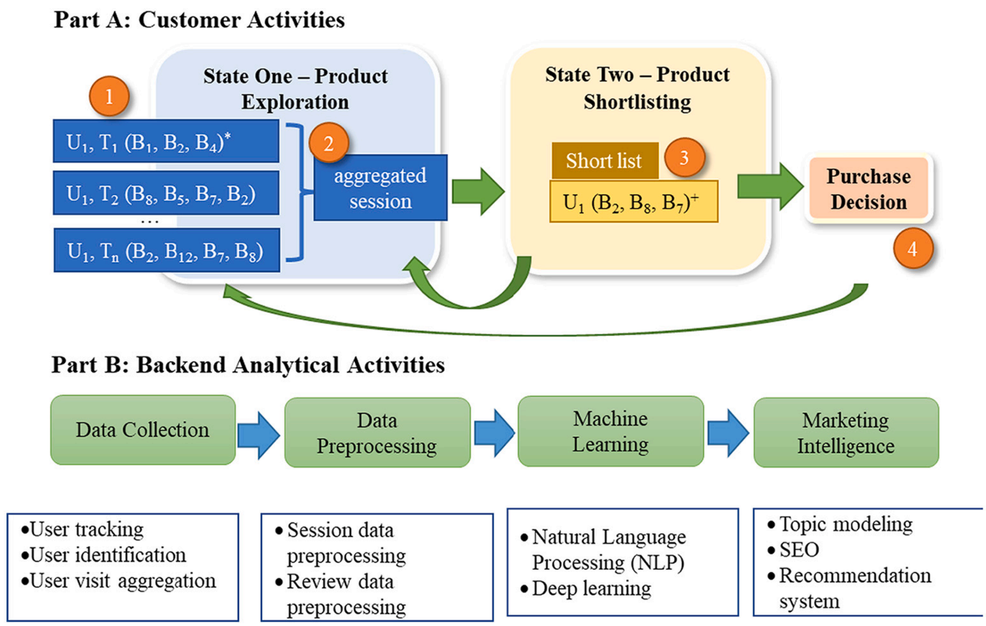


Fig. 1. Research framework.
* User 1 has reviewed pages of B&B hosts B_1, B_2 and B_4 during time 1 (T_1).
+ The shortlisted B&B hosts are B_2, B_8 and B_7 after analyzing user 1's (U_1) visit history.

multiple user sessions of the same individuals from different times. With these linked sessions, we are then able perform longitudinal tracking and extraction of user preferences from longer user visit history. Most importantly, the timing for the formation of steps two and three can therefore be reasonably calculated.

Part B of Fig. 2 is a flowchart of data collection and analysis as part of the novel marketing intelligence framework aimed to answer the research questions of the study. The whole process can be automated and steps are detailed in the sections below.

3.1. Data pre-processing

Users' long-term tracking sessions and related B&B reviews are key to this and later modules. Text mining techniques were performed to cleanse and extract the data before storing it in the database. B&Bs without any reviews were excluded because they do not offer UGC insights for users to narrow down their selections. Details of data preprocessing are outlined in the sub-sections below.

3.1.1. Session data preprocessing

Generally, there are two kinds of user sessions on web sites. Per-visit sessions (PVSs) are user sessions created by a web server at each visit. The three boxes under bubble one in Part A of Fig. 1 denotes three PVSs at times T_1, T_2 and T_3. The same individual visiting the web site at different times will be treated as different users and assigned different session IDs. Another type of sessions, termed long-term tracking sessions

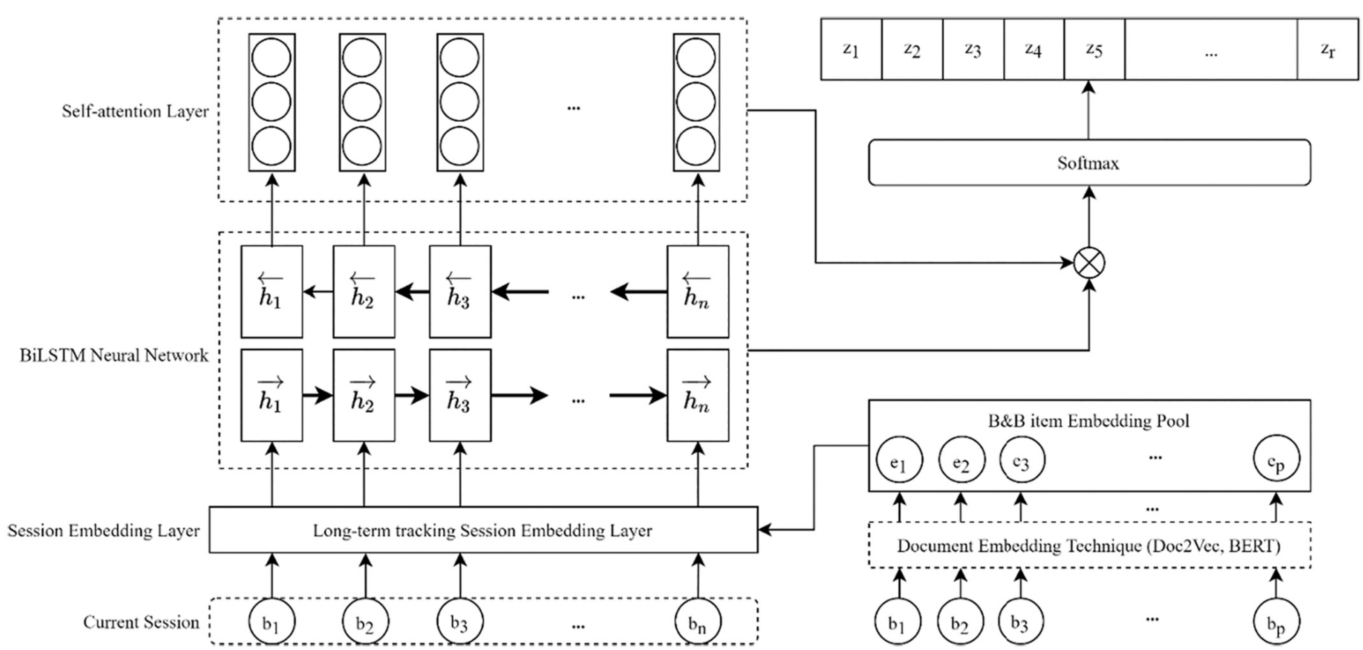


Fig. 2. The proposed recommendation framework.

(LTTSs), is hard to come by since it requires aggregation of individual PVSs for the same user. Bubble number two in Part A of Fig. 1 shows a LTTS for user 1 that aggregates the three PVSs visited at different time by the same user. Table 1 illustrates PVS and LTTS sessions. For example, the LTTS Id 12,345 includes three PVSs, one visited two B&B sites (see the bnb_id column) on 2/3/2018 and the others visited sites on 5/5/2018 and 5/9/2018.

The long-term tracking session records should be aggregated by per-visit session and noise (such as sessions not yet leading to a purchase) should be removed from the data. By doing so, we are then able to analyze the pattern of user visits for its effect on the final purchase.

3.1.2. Review data pre-processing

This step includes several text preprocessing and text mining techniques. All online reviews were collected and stored for each B&B as the basic unit for analysis. These review texts were then processed through pre-processing steps, including word segmentation, part-of-speech (POS) tagging, stop-word elimination, POS filtering, and keywords selection. Unlike the English language, the Chinese language does not use space to delimit words or phrases, which make it difficult to identify individual phrases. A popular specialized tool – Jieba [41], was used to extract Chinese word phrases. Key word phrases were screened by content experts for keyword clustering below.

3.2. Text mining

After preprocessing customer reviews, we followed the existing practice [25] to build term vectors from verbs, nouns and adjectives using Word2vec. The advantage of Word2vec is that words with similar meanings are close in their spatial positions. Words with similar meanings were aggregated through k-means cluster analysis to more precisely identify product attributes perceived by consumers. After using the elbow method to determine the best number of clusters, several candidate solutions were possible for the final number of clusters. We then consulted experts to review the cluster solution and label clusters that best represent the data.

3.3. Recommendation system

Key information at the B&B booking website relevant to this present study include B&B description, online reviews and user's browsing history. Browsing history is a collection of per-visit sessions and the B&B pages a user had visited since account creation. The process is represented in the following way. Let m denote the total number of possible B&Bs and B denotes the collection of all available B&Bs. $B = \{b_1, \dots b_m\}$, where b_j is the jth B&B in the collection B. In order to study the purchase behavior that is of interest to the practitioners, we analyzed only the long-term tracking sessions that eventually lead to a purchase. Given a set S of s long-term tracking sessions from p users, we have $S = \{s_1, \dots, s_p\}$. B&Bs visited in a long-term tracking session of user j is represented as an ordered list $s_j = [b_1, \dots b_q]$, where b_i is the ith B&B site visited by user j, and it satisfies $b_i \in B$.

As a recommendation system, its primary goal is to predict the next B&B site of most interest to the user (i.e. b_{q+1}). Let us define the number of candidates as r. Since there are many candidate B&Bs, a predicted result z_r through the model is represented as the probability of user's next click for r candidate B&Bs. Among r candidate B&Bs, top-k B&Bs with the highest probabilities will be recommended to user j. Further details are provided in Fig. 2 and the later sections.

3.3.1. B&B item embedding pool

All B&B′ reviews and host descriptions were captured and transformed into word vectors through a document embedding process before our deep learning analysis. For each B&B, we summarized all its reviews as a document and applied Doc2Vec and BERT document embedding techniques to transform it into a vector. In later experiments, we compared the performance of both and recommended the best algorithm for the task. The B&B item embedding pool contains the embedded document for each B&B independently computed from the review database. Therefore, the B&B item embedding pool can be seen as a repository of host descriptions and user reviews. B&B item embedding pool is denoted as $E = \{e1, \dots, e_m\}$, where e_i is the vectorized version of reviews and host description for the ith B&B. Each document vector e_i was a d dimensional vector, where d is the size of the vector document embedding.

3.3.2. User sessions and the session embedding layer

To better predict a user's next clicks, we combined B&Bs visited in the long-term tracking session and the most recent per-visit session to form an aggregate representation. The next step begins with feeding (a) the combined user sessions, and (b) product reviews and B&B host description from the B&B Item Embedding Pool, into the Session Embedding Layer that looks up B&B hosts visited by the user in a session. At this point, the user, visited B&B, and product reviews of those B&Bs are linked together in a vectorized form for internal processing. The input to this session embedding layer is the vectorized product reviews and host description of each visited B&B in a long-term tracking session $s_j = [b_1, \dots b_n]$. Each B&B id in the session was mapped to its id in the B&B item embedding pool and was converted into a d dimensional vector, where d is the size of the vector. Each B&B's review or host description was used by document embedding techniques as a d dimensional vector. The session embedding matrix S^{n*d} was constructed where n is the length of the session and d is the embedding size. The matrix is the input for the **BiLSTM Layer.**

3.3.3. The BiLSTM layer

The BiLSTM layer then performs a two-way bi-directional deep learning to learn from the above vectorized representation of user behavior. As an improvement over the traditional RNN, LSTM makes it easier to remember past data in memory and solves the vanishing gradient problem of RNN [31]. The advantage of bidirectional LSTM model (BiLSTM) is that it can learn user's long-term historical preference from the forward direction session data and the short-term preference from the backward direction. Additionally, a self-attention mechanism could be implemented to learn from repeat visits, a key aspect of consumer's shortlist. We combined BiLSTM and self-attention to distill the

Table 1
Example of user's two kinds of session data.

long-term tracking session _id	per-visit session _id	bnb_id	city	country	event_time	event_type
12,345	1,517,674,745,365	8639	KXX City	Taiwan	2018/2/3 16:22	Showed Host Page
12,345	1,517,674,745,365	725	KXX City	Taiwan	2018/2/3 16:31	Showed Host Page
12,345	1,525,516,806,694	1159	TXX City	Taiwan	2018/5/5 10:41	View - View Page
12,345	1,525,516,806,694	5061	TXX City	Taiwan	2018/5/5 10:42	View - View Page
12,345	1,525,516,806,694	9952	TXX City	Taiwan	2018/5/5 10:43	View - View Page
12,345	1,525,516,806,694	1461	TXX City	Taiwan	2018/5/5 10:58	View - View Page
12,345	1,525,871,596,664	1461	TXX City	Taiwan	2018/5/9 12:58	View - View Page
12,345	1,525,871,596,664	1159	TXX City	Taiwan	2018/5/9 13:14	View - View Page
12,345	1,525,871,596,664	1159	TXX City	Taiwan	2018/5/9 13:15	Complete-order

most important part of the user sequence. After that, the user session embedding matrix S^{n*d} is output as the hidden vector $\{h_1, h_2, \ldots, h_n\}$. The resulting output contains vectors with more than 100 dimensions, many of which contain zeros. Therefore, compaction is performed in the next layer.

3.3.4. Self-attention layer

The Self-Attention Layer compacts the input from the BiLSTM Layer to make the key user preferences stand out [42]. It uses RNN Encoder–Decoder to improve a machine translation in NLP. The main idea is to make the decoder decide the parts of the sentence to extract attention. Self-attention mechanisms have been found to improve attention models in the NLP literature [42].

Inspired by the idea of the attention mechanism, we used the self-attention mechanism for the hidden state vectors of bidirectional LSTM network to focus on the B&B's of interest to users. This makes frequently visited B&Bs stand out, which is analogues to consumer shortlisting their choices.

For each hidden vector of the BiLSTM neural network h_t was fed into the attention layer to calculate a new attention weight ∂_t. There were a set of hidden vector $\{h_1, h_2, \ldots, h_n\}$ and a set of attention weight$\{\partial_1, \partial_2, \ldots, \partial_n\}$.

For each user long-term tracking session, not all B&Bs had an equal contribution to represent user's preference. To address this issue, the self-attention mechanism is used to extract these more important B&Bs by giving them a higher weight ∂_t to increase their importance.

$$\widehat{s_j} = \sum_t \partial_t h_t \tag{1}$$

Finally, the weighted mean of the hidden vector h_t and attention weight ∂_t were combined to create a user's final preference session vector $\widehat{s_j}$ representing the embedded user session.

3.3.5. Softmax layer

The SoftMax Layer calculates the probability of user preference for each user and each candidate B&B recommendation. The output of self-attention layer was represented as a user's final preference session vector $\widehat{s_j}$. They are the r candidate B&Bs selected for recommendation prediction. For each candidate item B&B j's embedding are multiplied by a user's preference session vector to get a combined vector $\widehat{z_j}$ .Next, the softmax function calculates the probability that a user might be interested in a B&B. This is calculated for each user and each B&B. In the last neural network layer, the softmax function was used to normalize the output layer as a probability distribution, as follows:

$$\widehat{y} = softmax\left(\widehat{z_j}\right) \tag{2}$$

where $\widehat{y}$ the probability that a candidate B&B is likely to be clicked next in the long-term tracking session. The softmax function is a function that turns a vector of k real values into a vector of k probability values that add up to 1. In our framework, the probabilities produced by softmax represent the likelihood of purchase extracted from each user's long-term tracking session.

For each long-term tracking session, the loss function is defined as the categorical_crossentropy of the prediction and the ground truth. It can be written as follows:

$$\text{Loss} = -\sum_{i=1}^{m} y_i . log\widehat{y_i} \tag{3}$$

where y is a vector with the one-hot encoded label for each candidate B&B.

4. Experiments and analysis

The cleansed datasets include 43,031 browsing records and 44,166 customer reviews. There were 34,947 unique reviewers in the review dataset.

4.1. Priming and marketer defined primed messages

The objective of this section is to verify whether using the priming technique by B&B hosts is useful. After data preprocessing to transform unstructured B&B host descriptions into structured data, two popular word embedding techniques (Word2vec and Doc2vec) were used to represent words and documents as vectors for further machine learning processing.

4.1.1. Priming in a competitive environment

In this section, we report the outlook of the primed messages in B&B host descriptions on their product pages, since this is the first impression the B&B hosts had on potential consumers. Descriptions of all hosts that had received at least a confirmed purchase were collected and preprocessed before running Doc2vec embedding. We then performed pair comparisons of B&Bs within the same city on the similarity of their host descriptions. This generates a similarity score for each pair of B&B's in the same city. The maximum similarity score for each city is over 90%, meaning that there were at least two B&B's within each city that used very similar primed words in host descriptions. The average similarity scores of all pairs of B&Bs is almost 65%. After comparing similarity of a B&B host description with those of the top 10 popular B&Bs in the same city, the average similarity score calculated was about 65% as well. These numbers show that most B&B sites primed their host descriptions very similarly compared to the top selling B&Bs. These findings confirm our original hypothesis in that competition inspires imitation of the primed messages in host descriptions. Similarly primed words used in competitors' host descriptions will likely dilute the primed effect of each host, challenging the extent primed effects predicted by the priming theory.

4.1.2. Two-stage product evaluation behavior

We observed that 96% of users' long-term tracking sessions contained shortlisting behavior in that they started by visiting several B&B sites and gradually narrow it down to a few candidates (i.e., shortlists). They then visited B&Bs in this shortlist repeatedly. This shows that the majority of users follow the two-stage product evaluation phase. The remaining 4% who did not were mostly promotional sales where users went straight to purchase from a specific B&B with coupons or using vendor's promotions. Of those users who show the shortlisting behavior, 94% of them eventually purchase from the B&B on their shortlists. The average length of shortlists was 6 B&Bs shown as the red dashed line in Fig. 3. The y axis on right shows number of B&Bs in consumer shortlists.

Fig. 4 also shows the top 10 most popular B&Bs by city. The left y axis shows the number of B&Bs in a city whose host descriptions were at least 80% similar to the top 10 B&Bs in the city. This shows that host descriptions of B&Bs are very similar to the popular B&Bs from the same city. As a result of this similarity, primed effects at this stage (bubble one in Fig. 1) will likely be quite limited. One nice implication about this finding of a two-stage process is that the process of shortlist formation can now be captured. By automating the process of analyzing long-term tracking sessions and storing the previously learned results from these sessions, marketers are now able to pin-point the time of shortlist formation. Subsequent priming and promotion will now be possible before the user makes a purchase decision. Predictive machine learning models or recommender systems may be built to predict likely user preferences from the long-term tracking sessions and the shortlists to provide timely recommendations next time the user comes back to the platform (point four in Fig. 1).

4.1.3. Crafting the primed messages from UGC

In this section, we report findings regarding candidate primed keywords extracted from consumers, rather than from marketer crafted

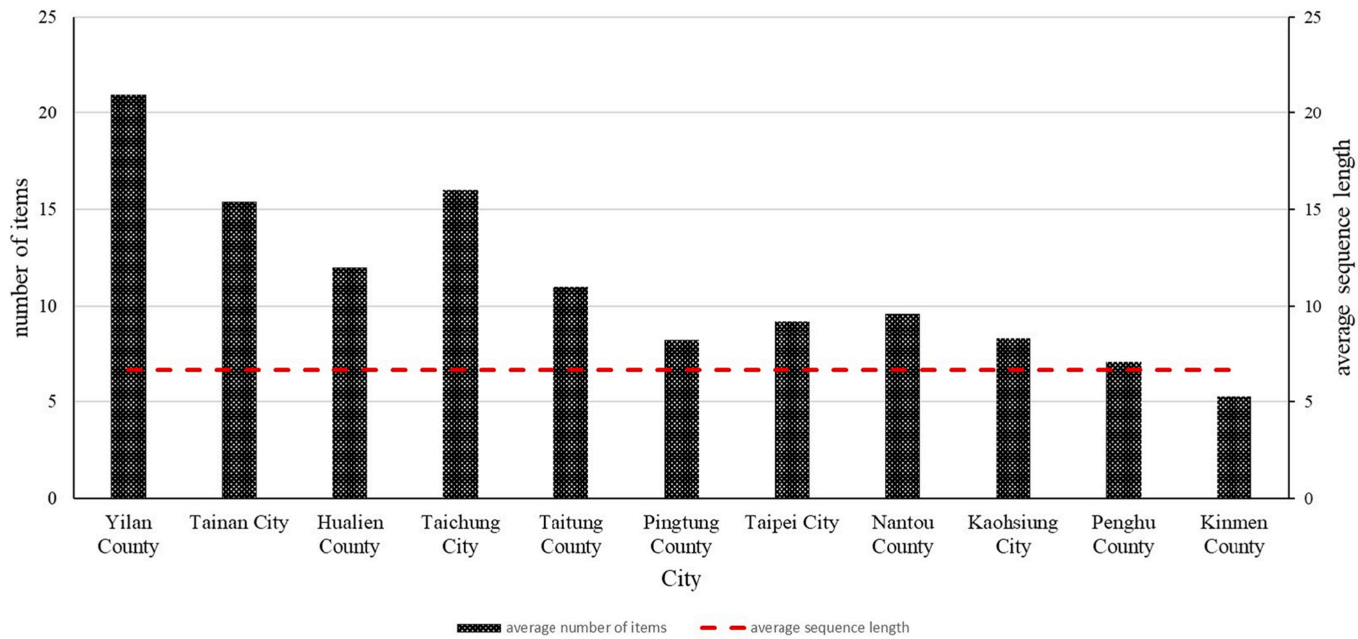


Fig. 3. Similarity of host descriptions.

Facilities	Decoration	Location	Hostel service
			
toilet (廁所), bathroom (浴室), sound proof walls (隔音), hope (希望), defects (缺點)	spaciousness (空間), comfort (舒服), warmth (溫馨), preference / liking (喜歡), impression / sensation (喜歡)	convenience (方便), traffic (交通), location (地點), night markets (夜市), surroundings (附近)	owner (老闆), friendliness (親切) service (服務), B&B (民宿)

Fig. 4. The result of words clouds.

primed messages. Instead of using a lab-controlled environment that involves a limited number of "potential" consumers, we opted to study user reviews. This provides two benefits. First, the results are more generalizable based on a large sample of actual user behavior. Second, using this large sample without coercing subjects is less prone to the Hawthorne Effect. In this experiment, we extracted verb, noun and adjective words from the corpus of 44,166 product reviews for clustering. After consulting with experts, we obtained 8 terms clusters. Four of which are illustrated in Fig. 4 to show the idea.

The first cluster contains customer complaints on facilities. Therefore, the cluster was named "facilities". The top five words include toilet (廁所), bathroom (浴室), sound proof walls (隔音), hope (希望), and defects (缺點). This cluster shows the areas the consumers wanted the B&B host to improve. The second cluster is "decoration". The top five words are spaciousness (空間), comfort (舒服), warmth (溫馨), preference / liking (喜歡), impression / sensation (喜歡). The adjectives used together with these nouns were positive words, such as warmth (溫馨), comfort (舒服), cuteness (可愛), and bright (明亮), all signifying the positives. The third cluster is related to location. The top five keywords are convenience (方便), traffic (交通), location (地點), distance to nearby night markets (夜市), and surroundings (附近). In this cluster most words are about the proximity of hotel to nearby attractions or bus stations. The fourth cluster is about hostel service. The hot words are host's hospitality (老闆), friendliness (親切) and service (服務) with positive adjectives.

As these words and clusters are generated from user reviews, they offer several key benefits for the B&B hosts. First, the words reported in Fig. 5 could be used as the priming stimuli, and the words together with the priming stimuli in the same cluster could be viewed as the words that are activated by the priming stimuli. Therefore, using those words and word associations is a strategic tool for markers to craft the right primed messages better tuned based on user reviews. Second, the word clusters in this section could also be useful for SEOs and other marketing purposes.

4.2. Recommendation evaluation result

This section reports an experiment testing the efficiency of our recommendation system with several deep learning models. The goal is to empirically determine the best recommendation model capable of predicting the user's next click. With this best model, it could be used to determine the best timing to continue subsequent priming or promotion in steps two and three from Fig. 2.

We used document embedding techniques (Doc2Vec and BERT) to represent each B&B reviews and host description as word vectors. We proposed and compared the performance of three session-based recommendation models, including bidirectional LSTM with self-attention, bidirectional LSTM and bidirectional GRU. A total of 30,688 long term tracking sessions were collected for this study and they were used for the comparison of the above embedding techniques. Long-term tracking sessions considered in the present study are those that lead to a purchase. We used 5-fold cross validation to check robustness of the resulting models. Data was split into 60% training, 20% validation, and 20% test. Results are presented in Table 2. Hit rate, HR@k, was used as the performance evaluation metric, which is widely used in previous session-based evaluations [31,32,38]. HR@k is the likelihood the next consumer click is among the top-k B&Bs recommended by the system. GRU4rec [43] was used as the baseline for performance comparison of session-based recommendation systems (see [31]) and variants were compared below:

- ➢ **(Doc2Vec + BiLSTM) model**: The model applied Doc2Vec model for the document embedding and LSTM model for the session embedding architecture.
- ➢ **(Doc2Vec + GRU) model**: A variant of previous model which applied GRU model for the session embedding architecture.
- ➢ **Doc2Vec + BiLSTM (self-attention) model:** A variant of previous model which only applied LSTM with attention mechanism for the session embedding architecture.
- ➢ **(BERT+ BiLSTM) model**: The model applied BERT model for the document embedding and LSTM model for the session embedding architecture.
- ➢ **(BERT + GRU) model**: A variant of previous model which applied GRU model for the session embedding architecture.
- ➢ **BERT+ BiLSTM (self-attention) model**: A variant of previous model which only applied LSTM with attention mechanism for the session embedding architecture.

Table 2
The accuracy result of the proposed models using review embedding and host embedding.

	B&B review			Host Description		
	HR@5	HR@10	HR@20	HR@5	HR@10	HR@20
Baseline (GRU4Rec)	0.558	0.629	0.713	0.558	0.629	0.713
Doc2Vec + CNN+ DNN	0.461	0.582	0.632	0.502	0.590	0.666
Doc2Vec + CNN + RNN	0.650	0.705	0.755	0.601	0.664	0.721
Doc2Vec + CNN + BiLSTM	0.748	0.804	0.837	0.714	0.767	0.809
Doc2Vec + BiLSTM	0.885	0.902	0.918	0.870	0.894	0.911
Doc2Vec + GRU	0.826	0.849	0.875	0.803	0.833	0.857
Doc2Vec + BiLSTM (self-attention)	**0.886**	**0.903**	**0.918**	**0.879**	**0.897**	**0.918**
BERT+ BiLSTM	0.775	0.811	0.844	0.818	0.845	0.870
BERT+GRU	0.7582	0.794	0.829	0.767	0.803	0.834
BERT + BiLSTM (self-attention)	**0.866**	**0.888**	**0.908**	**0.871**	**0.898**	**0.918**
BERT + BiLSTM (self-attention) with aspect term	**0.9073**	**0.9247**	**0.9434**			

The overall performance in terms HR@5, HR@10, HR@20 is shown in Table 2 with the best results highlighted in boldface. To prevent overfitting, the dropout rate is set to 0. The number of layers is 11. The batch size is 16 and the optimization algo is adam.

Table 2 shows that BiLSTM proposed in our model greatly improved the model accuracy over the baseline model. With self-attention added, the model accuracy is further enhanced for both Doc2Vec and BERT (see the first and second bolded rows in Table 2). Although both Doc2Vec and BERT are robust, BERT is generally less accurate compared to Doc2Vec. This is because of the limitation of the underlying algorithm of BERT that considers only the first 512 words (see [44]). Compared with Doc2Vec that considers all reviews despite their lengths, this word limit of BERT only allows for capturing the first few reviews. As a result, its accuracy is affected. Here we propose an enhancement that aggregates only the parts of reviews containing those keywords (i.e., aspect terms) identified in section 4.1.3, as opposed to aggregating the whole review messages. The last row of Table 2 reflects the result of this enhancement, which even makes BERT to outperform Doc2Vec. Since the generation of those priming keywords in section 4.1.3 can be done in a batch manner infrequently, doing so does not greatly increase the amount of time needed to process BERT. This enhancement through priming keywords is usually not needed for host description because a host description is generally within the 512 word limit. Thus, cells under host description for the last row of Table 2 are empty. The accuracy figures for host descriptions are the same as those in the second bolded row.

Moreover, we conducted the effect size test using Cliff's delta [45] to show the statistical significance of our experimental results. Table 3 shows the results of the test across the models. When Cliff's Delta is 1, it indicates that the effect size of one model is larger than the other model. As Table 3 shows, the proposed models Doc2Vec + BiLSTM (self-attention) and BERT+BiLSTM (self-attention) are superior to other models.

4.3. Alternative view of the proposed models

As we have seen in the previous section, BiLSTM and self-attention added to Doc2Vec and BERT enhanced their performance. Adding aspect terms further improved BERT's performance. Here in this section we present two additional views on the proposed algorithms to show their robustness, one that concerns about search vs. experience terms, and the other one that concerns about similarity vs. dis-similarity of product reviews.

Table 3
The Cliff's delta effect size test results of models.

Models		HR@5
Doc2Vec + BiLSTM (self-attention)	BERT+BiLSTM (self-attention)	0.22
Doc2Vec + BiLSTM (self-attention)	Baseline	1
BERT+BiLSTM (self-attention)	Baseline	1
Doc2Vec + BiLSTM (self-attention)	Doc2Vec + GRU	1
BERT+BiLSTM (self-attention)	BERT+GRU	1

In the first view, BnB reviews and host descriptions were split into search term dominant versus experience term dominant groups. A review is classified as experience term dominant if it includes mostly the information regarding keywords from the Decoration and Hostel Service clusters identified in section 4.1.3. A review is classified as search term dominant if it includes mostly the keywords from the Facility and Location clusters. The distinction between the two is analogues to search goods versus experienced goods, where product quality of search goods can be gauged by simply searching and estimating from the user's collected information, while product quality of experienced goods is approximated by learning from others' reported experiences. The first two rows in Table 4 report model performance on the two groups respectively. As the accuracy figures show, both Doc2Vec and BERT with our proposed enhancements were still quite robust with the minimum accuracy being 0.908 for product reviews and 0.911 for host descriptions.

The second view is inspired by another line of study that concerns about possible social influence on review writing (see [46–48]). Although an accurate measure of social influence among review writers is evolving, the end result of social influence is similar reviews being produced (see [46–49]). To gauge review similarity, we first built a vector of keyword occurrences for each review based on the dimensions defined in section 4.1.3. Reviews for the same BnB that are similar in their keyword vectors were put into the Similarity group, while reviews that are different were listed in the Dis-similarity group. As the second half of Table 4 shows, both Doc2Vec and BERT with our proposed enhancements were still robust with the lowest accuracy for the Similarity group being 0.911 for product reviews and 0.90 for host descriptions. The lowest accuracy figures for the Dis-similarity group were 0.941 for the product reviews and 0.913 for host descriptions. Both alternative views have demonstrated that our proposed enhancements are robust across different data sets and different levels of granularity.

5. Discussion and conclusion

Priming is popular in marketing and advertising, but its effect is challenged by competing messages in an online environment and by eWoM (online product reviews). Competition in an online environment dilutes an originally primed message rendering it less effective for the intended purpose, while eWoM is another powerful source to sway potential customers. Priming also loses its effect after consumers shortlisting products. Products selected into the shortlist present similar messages with similar appeals. The heuristics of the final selection for purchase is usually internal to individual consumers. As this shortlist may expand and reduce, messages that relies on the rigidly defined primes from the mainstream methodology (lab experiments and surveys) can easily lose their applicability. These issues severely limit the applicability of the priming theory. Therefore, we tested a framework to approximate the user heuristics after shortlisting using NLP and deep learning to capture the information consumers specifically visited during their long and short-term user sessions.

To answer the above questions, we first propose a two-stage product evaluation process and tested it through a series of NLP, and deep learning techniques to provide a novel approach for marketing intelligence. The first stage constitutes one of more visits to B&B sites screening their offerings before shortlisting. Stage two is the actual product shortlisting stage before the final purchase. The two product evaluation stages involve four steps illustrated in Part A of Fig. 1. Steps one and four are typically visible to marketers where priming can be performed. The challenge faced in these steps is to find the right primed words and their associated network of words activated by them. Steps two and three are more difficult for traditional approaches since shortlisting is usually internal to users and the shortlist does not always stay static. In the present study, we address these challenges with the following strategies. First, clustering of keywords extracted from product review data visited repeatedly after shortlisting provides ample insights into how certain words or topics that are likely to be mentioned together. This serves as the candidate priming keywords and their associated network of activated words. Second, our approach of tapping into consumers' revisit patterns in long-term tracking sessions allows us to pinpoint the formation of shortlists, thereby making a wide range of opportunities (e.g., subsequent priming, promotion and differential fees) possible during steps two and three in Fig. 2. There are few studies about user long-term tracking session because users' preference and historical records are hard to obtain. We proposed a novel session-based recommendation system that incorporates RNN models into representing long-term tracking sessions. The proposed deep learning model based on the bidirectional LSTM technique with the self-attention technique mimics user' two-stage decision process. Multiple recommendation systems were compared to identify the optimal combination of hybrid recommendation systems.

5.1. Theoretical and methodological contributions

Through a comprehensive literature survey on NLP studies, Kang & Cai [18] pointed out that it was difficult to make theoretical contributions purely by NLP alone because their results are not linked to existing management theories. In this present study, we attempt to fill this gap by linking NLP and deep learning to the popular priming and eWOM theories. More specifically we identified where the priming theory falls short, and how NLP and deep learning could be used to capture consumer heuristics after shortlisting. As a result, the priming theory can therefore be extended into this shortlisting phase that was not possible with the traditional means before. This answers our RQ1. While capturing the user heuristics with deep learning, we also compared single directional deep learning techniques (GRU and traditional RNN) with bidirectional LSTM, as well as with and without the attention mechanism added. The findings offer the optimized version of deep

Table 4
Alternative views of algorithm robustness.

	Review						Description					
	Search terms			Experience terms			Search terms			Experience terms		
	HR @ 5	HR @ 10	HR @ 20	HR @ 5	HR @ 10	HR @ 20	HR @ 5	HR @ 10	HR @ 20	HR @ 5	HR @ 10	HR @ 20
Doc2Vec + LSTM (self-attention)	0.944	0.958	0.965	0.952	0.964	0.971	0.937	0.951	0.9639	0.944	0.954	0.962
BERT + LSTM (self-attention) with Aspect terms	0.940	0.947	0.958	0.908	0.928	0.943	0.911	0.944	0.956	0.937	0.949	0.961
	Similarity			Dis-similarity			Similarity			Dis-similarity		
Doc2Vec + LSTM (self-attention)	0.924	0.937	0.949	0.946	0.955	0.9678	0.90	0.915	0.941	0.913	0.933	0.953
BERT + LSTM (self-attention) with Aspect terms	0.911	0.924	0.932	0.932	0.9415	0.9547	0.905	0.920	0.949	0.926	0.946	0.957

learning models best for mining both long- and short-term sessions. This answers our RQ2.

Our work presents several theoretical contributions. First, the primed effect weakens as the time goes by [9,50], and especially when a consumer starts consulting a different source of information (e.g., eWOM) [51]. Our two-stage framework bridges the priming theory and the eWOM, to provide a more realistic theoretical extension of existing theories. Second, the combination of NLP and deep learning techniques in the present study captures consumer heuristics after shortlisting that is internal to individual consumers, which offers marketers opportunities to perform additional priming or timely promotion during the time (e.g., steps 2 and 3 in Fig. 2) that is difficult for traditional means. By analyzing traces of user activities after shortlisting, we not only extend the "shelf-life" of priming theory into the shortlisting phase, but also make it possible for subsequent priming to be performed closer to actual purchase – a effect known to have a greater success on affecting purchase [52]. Third, our work also represents a methodology improvement by simultaneously considering both long and short-term user visit sessions, and optimizing the deep learning representation of these extracted user sessions.

5.2. Managerial implications

Our work answers several challenges of marketing theory and decisions with a framework consisting of optimized deep learning models tested across multiple data sets. As a result, the findings are of relevance to both technical (IS/IT and data science) and marketing managers. Models built from the proposed framework provide several technical implications. First, as shown in the present study, aggregation of long- and short-term user sessions offers a useful approximation of user heuristics used after shortlisting. Short-term sessions represent a consumer's immediate needs, while the long-term sessions capture a consumer's innate preference. By considering both, our framework is closer in modeling user behavior after shortlisting. The key to associate multiple discrete user sessions is the ability to identify each individual user before aggregating their user sessions. Literature shows several promising techniques ranging from profiling, behavioral-matching, and hardware & software identifications [53] that may be considered. Multiple of these techniques may also be used together to increase accuracy of user identification.

Second, despite that the two deep learning techniques were shown in our study to be robust in varied data sets. BERT is susceptible to its 512-word limits [44]. Such a limitation becomes a challenge when the aggregated data exceeds this limit. In our case of BnB reviews, this algorithmic limitation allowed only the first few reviews from each BnB to be analyzed after aggregating all reviews. This hampers BERT's accuracy. Our approach of distilling keyword clusters (section 4.1.3) and using parts of reviews containing those keywords before running BERT further enhanced the algorithm's accuracy. Since the keyword distilling process does not have to be run in real-time or frequently, it adds only marginal overhead to the algorithm.

Third, our framework shows that a mere implementation of deep learning models would not automatically produce an accurate model. Doc2Vec and BERT in their original forms were only accurate at about 40% - 60% as shown in Table 2. With BiLSTM and self-attention added, the accuracy is greatly improved. The reason of why this addition produces a better result is because it takes repeat visits into consideration. This repeat visit not only shows user interest in a BnB, but also serves as the basis to identify the formation of shortlists.

Because of the above technical advancements in our tested models, the following opportunities becomes available for marketing managers:

First, our framework allows monitoring the formation of shortlists and products selected. With this knowledge, attributes such as product similarity (e.g., price and features), primed keywords used by competitors, and patterns of repeat visits can now be captured for the type of marketing decision-making that was difficult before. This knowledge also helps identify the right timing for subsequent priming and further promotions (e.g., instant coupons) before purchase. Our approach if automated allows the flexibility to work with shortlists that does not always stay static. Practitioners should consider adopting our optimized deep learning models (Doc2Vec + BiLSTM with attention and BERT+BiLSTM with attention) to best capture consumer heuristics.

Second, our work identified clusters of keywords congruent to user preferences as keywords were extracted from user reviews that consumers repeatedly visited. They give us a glimpse into the related words activated in user's minds – an area emphasized in the priming theory, but is difficult to identify in practice due to these nodes or words being internal to the users. The value of our contribution here is that not only we present candidate priming keywords extracted from consumers to model their use heuristics, we also present words in the same cluster together that approximates the network of words activated by the primed keywords as prescribed by the priming theory. This latter part is very significant since word activation is internal to individual consumers. It is difficult to tap into this internal network of activated words with the existing approaches. Using this network of activated keywords in SEOs, promotional materials and subsequent priming will create a better association of the primed messages with the brand. Since features specific to individual products are valued in the product screening stage [5], practitioners may want to consider the keywords in product feature clusters (e.g., decoration and convenience clusters). Similarly, experience-specific product attributes are key in the shortlisting stage [5]. Experience specific keywords (e.g., keywords in the hostel service dimension) are relevant in this stage.

Our work extends the boundaries of the priming theory, enhances deep learning algorithm and enables marketing decisions during shortlisting. It also provides initial evidence to address managerial and theoretical contributions with inductive methodologies (machine learning deep learning, NLP, etc.). It is an initial step to respond to Kang & Cai's [18] call for research to bridge machine learning and management theories to produce new insights.

5.3. Limitations and future research

Researchers, business analysts, and practitioners could adopt the framework tested in our study to advance their marketing decision-making on priming, SEO and competitive analysis after shortlisting. However, our study is not without limitations. First, despite that our work extends the priming theory into the shortlisting phase, readers are cautioned not to extend our findings to non-experience goods (e.g., commodity and named brands of electronics) without testing them. This is because criteria used to review these non-experienced goods rely more heavily on attributes and pricing rather than usage experience of peers [6]. Second, word clustering may be extended with deep leaning methods for topic modeling to answer additional insights. Additionally, future research could build on our work to study how social influence [47] and review manipulation techniques influence how consumers perceive them and how that affects the boundary of the priming theory.

Data availability

The authors do not have permission to share data.

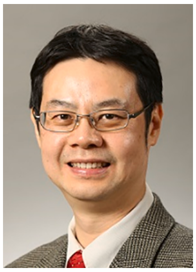
Acknowledgements

The authors are very grateful to the anonymous referees and Editor for their helpful comments and valuable suggestions for improving the earlier version of the paper. The authors would like to thank Cheng-Ting Chao and Chao-Min Hou for their assistance in processing data. This study was supported in part by the Ministry of Science and Technology of Taiwan under grants MOST 109-2218-E-002-016, MOST 108-2410-H-027-020 and MOST 109-2410-H-027 -009 -MY2.

References

[1] P. Borah, Seeking more information and conversations: influence of competitive frames and motivated processing, Commun. Res. 38 (3) (2011) 303–325.
[2] J.A. Bargh, P.M. Gollwitzer, A. Lee-Chai, K. Barndollar, R. Trötschel, The automated will: nonconscious activation and pursuit of behavioral goals, J. Pers. Soc. Psychol. 81 (6) (2001) 1014.
[3] U.R. Orth, H.F. Koenig, Z. Firbasova, Cross-national differences in consumer response to the framing of advertising messages: an exploratory comparison from Central Europe, Eur. J. Mark. 41 (3/4) (2007) 327–348.
[4] S. Tanford, M. Kim, E.J. Kim, Priming social media and framing cause-related marketing to promote sustainable hotel choice, J. Sustain. Tour. 28 (11) (2020) 1762–1781.
[5] A.R. Dennis, L. Yuan, X. Feng, E. Webb, C.J. Hsieh, Digital nudging: numeric and semantic priming in e-commerce, J. Manag. Inf. Syst. 37 (1) (2020) 39–65.
[6] P. Jones, M.-M. Chen, Factors determining hotel selection: online behaviour by leisure travellers, Tour. Hosp. Res. 11 (1) (2011) 83–95.
[7] D.J. Zhang, et al., How do price promotions affect customer behavior on retailing platforms? Evidence from a large randomized experiment on Alibaba, Prod. Oper. Manag. 27 (12) (2018) 2343–2345.
[8] H.C. Murphy, M.-M. Chen, M. Cossutta, An investigation of multiple devices and information sources used in the hotel booking process, Tour. Manag. 52 (2016) 44–51.
[9] J. DeCoster, H.M. Claypool, A meta-analysis of priming effects on impression formation supporting a general model of informational biases, Personal. Soc. Psychol. Rev. 8 (1) (2004) 2–27.
[10] E.A. Minton, T.B. Cornwell, L.R. Kahle, A theoretical review of consumer priming: prospective theory, retrospective theory, and the affective–behavioral–cognitive model, J. Consum. Behav. 16 (4) (2017) 309–321.
[11] M.R. Quillian, Word concepts: a theory and simulation of some basic semantic capabilities, Behav. Sci. 12 (5) (1967) 410–430.
[12] M. Posner, C. Snyder, R. Solso, Information Processing and Cognition: The Loyola Symposium, 1975.
[13] R.K. Minas, A.R. Dennis, Visual background music: creativity support systems with priming, J. Manag. Inf. Syst. 36 (1) (2019) 230–258.
[14] E.C. Nisbet, P.S. Hart, T. Myers, M. Ellithorpe, Attitude change in competitive framing environments? Open−/closed-mindedness, framing effects, and climate change, J. Commun. 63 (4) (2013) 766–785.
[15] R. Spiegler, Competitive framing, Am. Econ. J. Microecon. 6 (3) (2014) 35–58.
[16] M. Etco, S. Sénécal, P.-M. Léger, M. Fredette, The influence of online search behavior on consumers' decision-making heuristics, J. Comput. Inf. Syst. 57 (4) (2017) 344–352.
[17] X. Lu, S. He, S. Lian, S. Ba, J. Wu, Is user-generated content always helpful? The effects of online forum browsing on consumers' travel purchase decisions, Decis. Support. Syst. 137 (2020), 113368.
[18] Y. Kang, Z. Cai, C.-W. Tan, Q. Huang, H. Liu, Natural language processing (NLP) in management research: a literature review, J. Manag. Analyt. 7 (2) (2020) 139–172.
[19] Y. Wu, E.W. Ngai, P. Wu, C. Wu, Fake online reviews: literature review, synthesis, and directions for future research, Decis. Support. Syst. 132 (2020), 113280.
[20] L. Zhang, L. Hanks, Online reviews: the effect of cosmopolitanism, incidental similarity, and dispersion on consumer attitudes toward ethnic restaurants, Int. J. Hosp. Manag. 68 (2018) 115–123.
[21] Y. Guo, S.J. Barnes, Q. Jia, Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation, Tour. Manag. 59 (2017) 467–483.
[22] S. Gao, O. Tang, H. Wang, P. Yin, Identifying competitors through comparative relation mining of online reviews in the restaurant industry, Int. J. Hosp. Manag. 71 (2018) 19–32.
[23] A. Kumar, R.D. Gopal, R. Shankar, K.H. Tan, Fraudulent review detection model focusing on emotional expressions and explicit aspects: investigating the potential of feature engineering, Decis. Support. Syst. 155 (2022), 113728.
[24] B. Biswas, A. Mukhopadhyay, S. Bhattacharjee, A. Kumar, D. Delen, A text-mining based cyber-risk assessment and mitigation framework for critical analysis of online hacker forums, Decis. Support. Syst. 152 (2022), 113651.
[25] Y. Jung, Y. Suh, Mining the voice of employees: a text mining approach to identifying and analyzing job satisfaction factors from online employee reviews, Decis. Support. Syst. 123 (2019), 113074.
[26] Y.-H. Hu, K. Chen, Predicting hotel review helpfulness: the impact of review visibility, and interaction between hotel stars and review ratings, Int. J. Inf. Manag. 36 (6) (2016) 929–944.
[27] M. Jawad, R. Benbunan-Fich, Impact of information presentation format on user decision-making: a format-stage fit perspective, AMCIS 2020 Proceedings. 5 (2020).
[28] R. Filieri, E. Raguseo, C. Vitari, When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type, Comput. Hum. Behav. 88 (2018) 134–142.
[29] B. Hazarika, K. Chen, M. Razi, Are numeric ratings true representations of reviews? A study of inconsistency between reviews and ratings, Int. J. Bus. Informat. Syst. 38 (1) (2021) 85–106.
[30] A. Lawani, M.R. Reed, T. Mark, Y. Zheng, Reviews and price on online platforms: evidence from sentiment analysis of Airbnb reviews in Boston, Reg. Sci. Urban Econ. 75 (2019) 22–34.
[31] M. Ludewig, D. Jannach, Evaluation of session-based recommendation algorithms, User Model. User-Adap. Inter. 28 (4–5) (2018) 331–390.
[32] S. Wang, L. Cao, Y. Wang, Q.Z. Sheng, M. Orgun, D. Lian, A Survey on Session-Based Recommender Systems, arXiv preprint arXiv:1902.04864, 2019.
[33] S. Zhang, L. Yao, A. Sun, Y. Tay, Deep learning based recommender system: a survey and new perspectives, ACM Comput. Surveys (CSUR) 52 (1) (2019) 1–38.
[34] B. Hidasi, A. Karatzoglou, Recurrent neural networks with top-k gains for session-based recommendations, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 843–852.
[35] R. Mishra, P. Kumar, B. Bhasker, A web recommendation system considering sequential information, Decis. Support. Syst. 75 (2015) 1–10.
[36] M.K. Najafabadi, A.H. Mohamed, M.N.R. Mahrin, A survey on data mining techniques in recommender systems, Soft. Comput. 23 (2) (2019) 627–654.
[37] R. Mehta, K. Rana, A review on matrix factorization techniques in recommender systems, in: 2017 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA), IEEE, 2017, pp. 269–274.
[38] S. Wang, L. Cao, Y. Wang, Q.Z. Sheng, M.A. Orgun, D. Lian, A survey on session-based recommender systems, ACM Comput. Surveys (CSUR) 54 (7) (2021) 1–38.
[39] K. Cho, et al., Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, arXiv preprint arXiv:1406.1078, 2014.
[40] J. Li, P. Ren, Z. Chen, Z. Ren, T. Lian, J. Ma, Neural attentive session-based recommendation, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017, pp. 1419–1428.
[41] Jieba Chinese Text Segmentation. https://github.com/fxsjy/jieba accessed.
[42] A. Vaswani, et al., Attention is all you need, in: Advances in Neural Information Processing Systems, 2017, pp. 5998–6008.
[43] B. Hidasi, A. Karatzoglou, L. Baltrunas, D. Tikk, Session-based Recommendations with Recurrent Neural Networks, arXiv preprint arXiv:1511.06939, 2015.
[44] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805, 2018.
[45] N. Cliff, Dominance statistics: ordinal analyses to answer ordinal questions, Psychol. Bull. 114 (3) (1993) 494.
[46] A.K. Jha, S. Shah, Disconfirmation effect on online review credibility: an experimental analysis, Decis. Support. Syst. 145 (2021), 113519.
[47] S. Zhou, B. Guo, The order effect on online review helpfulness: a social influence perspective, Decis. Support. Syst. 93 (2017) 77–87.
[48] H. Li, F. Meng, M. Jeong, Z. Zhang, To follow others or be yourself? Social influence in online restaurant reviews, Int. J. Contemp. Hosp. Manag. 32 (3) (2020) 1067–1087.
[49] S. Sridhar, R. Srinivasan, Social influence effects in online product ratings, J. Mark. 76 (5) (2012) 70–88.
[50] L.R. Squire, A.P. Shimamura, P. Graft, Strength and duration of priming effects in normal subjects and amnesic patients, Neuropsychologia 25 (1) (1987) 195–210.
[51] J.H. Byrne, Learning and Memory: A Comprehensive Reference, Academic Press, 2017.
[52] A. Krishna, M. Wagner, C. Yoon, R. Adaval, Effects of extreme-priced products on consumer reservation prices, J. Consum. Psychol. 16 (2) (2006) 176–190.
[53] Y.C. Yang, Web user behavioral profiling for user identification, Decis. Support. Syst. 49 (3) (2010) 261–271.

Dr. Li-Chen Cheng is currently a Professor of Department of Information and Finance Management, National Taipei University of Technology, Taipei, Taiwan. She received her Ph.D. degree in information management from National Central University. Dr Cheng serves as an editorial board member and a reviewer for more than 10 academic journals. She is an associate editor of Electronic Commerce Research and Applications now. Her research interests include deep learning, opinion mining, financial technique, AI in internet marketing, business intelligence and decision making models. She has published papers in in well-recognized SSCI and SCI journals including Decision Sciences, Decision Support Systems, Information Processing & Management, Electronic Commerce Research and Applications, Journal of Information Science, Neurocomputing, European Journal of Operational Research, and many others.



Dr. Kuanchin Chen is Professor of Computer Information Systems, Director of the Center for Business Analytics, and John W. Snyder Fellow at Western Michigan University. He has more than twenty years of research, teaching and consulting experience. His research interests include electronic business, data analytics, social networking, privacy & security, online behavioral issues, machine learning and human computer interactions. Dr. Chen has published in many reputable academic journals, including Information Systems Journal, Decision Support Systems, Information & Management, IEEE Transactions on Systems, Man, and Cybernetics, International Journal of Information Management, Journal of Database Management, Internet Research, Communications of the Association for Information Systems, Electronic Commerce Research and Applications, Journal of Global Information Management, DATA BASE for Advances in Information Systems, Information processing & Management, International Journal of Medical Informatics, and Tourism Management. He has been an editor, associate editor, and editorial member of several scholarly journals. He is also a recipient of several research and teaching awards, including awards given by scholarly journals & conferences, department, college, university and U.S. Fulbright program. He is frequently invited to give research talks at universities, government agencies and other institutions.