Information & Management 59 (2022) 103719

Contents lists available at ScienceDirect

Information & Management

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

Your posts betray you: Detecting influencer-generated sponsored posts by finding the right clues

Rong-Ping Shen [a,b], Dun Liu [*,a,b], Xuan Wei [c], Mingyue Zhang [d]

[a] School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China
[b] Key Laboratory of Service Science and Innovation of Sichuan Province, Southwest Jiaotong University, Chengdu 610031, China
[c] Antai College of Economics and Management, Shanghai Jiaotong University, Shanghai 200030, China
[d] School of Business and Management, Shanghai International Studies University, Shanghai 201620, China

ARTICLE INFO

Keywords:
eWOM
Influencer marketing
Sponsored content detection
Deep learning
Task-specific clues

ABSTRACT

With the prevalence of sponsorship practice using social media posts, the detection of sponsored content becomes crucial for platforms to regulate the generated content and prevent users from being misinformed. However, there is a paucity of investigations on the detection of sponsored content in existing research. To fill this research gap, we first identify several task-related clues by referring to relevant psychological theories and practical observations. Based on the clues, we conceptualize four types of sponsored content features and propose a unified deep learning detection framework, which also learns the relative importance of each feature. Experiments conducted on 26,823 social media posts demonstrate the performance of our proposed model compared with competitive alternatives and the value of each feature. The learned feature importance also enables deeper phenomena understanding. The research findings provide actionable insights into the narrative strategies influencers adopt and how to distinguish online sponsored content.

1. Introduction

The exponential development of social media has widened the reach of electronic word-of-mouth (eWOM) effect and deeply innovated the formats of marketing industry [1,2]. Those who have created attractive content and thus garnered substantial followers on social media are commonly identified as social media influencers (SMIs). The core of SMI is rooted in their potential influence, which enables SMIs to disseminate opinions and shape the attitudes of their followers through creating content on channels such as Instagram, YouTube, and Twitter [3,4]. Brand marketers have long recognized the potential impact of SMIs and parlayed the influencer-follower relationship into a wide variety of marketing practice. One of the most prevalent influencer marketing formats is sponsored posts, that is, SMIs receive diverse types of direct/indirect monetary compensation and in turn incorporate brand-related content (brand endorsements or product mentions) in their posts to help broadcast social buzz around brands. Comparing with traditional marketing techniques (e.g., celebrity endorsement), influencer marketing outweighs in being perceived as intrinsically motivated and non-commercial orientation driven. Additionally, comparing with other formats of sponsorship practice (e.g., sponsored review), the impact scope of sponsored posts is greatly widened and underpinned by massive social broadcast from influencers' solid follower base [5,6].

While influencer marketing is proven to bring a boost in sales volumes, its commercial nature also inevitably shows and breeds a trust crisis from consumers [7–9]. On social media platforms, another voice of eWOM with real-life experience and objective opinions, defined as organic posts, also forms a powerful force. With the accumulated feedback from sponsored posts, the concealment of sponsorship began to be exposed by a contradictory voice from authentic consumers [10,11]. As a consequence, SMIs, brands, and even the platform itself suffered from delayed but more severe criticism as biased, unreliable, and even deceptive. In this regard, how to automatically distinguish sponsored content becomes critical for platform managers to effectively regulate platform content and prevent potential reputation damage.

With the influx of sponsored content into all types of e-commerce websites and social media channels, platform managers have tried several measures to regulate its practice. For example, Amazon required reviewers on its platform to include a sponsorship disclosure statement if the reviewer has received any types of compensation. However, in most cases, users are reluctant to voluntarily disclose sponsorship despite that platforms have warned to conduct strict punishments towards dishonest

* Corresponding author at: School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China.
E-mail address: newton83@163.com (D. Liu).

https://doi.org/10.1016/j.im.2022.103719
Received 19 January 2022; Received in revised form 27 September 2022; Accepted 5 October 2022
Available online 8 October 2022
0378-7206/

behaviors. As a consequence, some platforms have to hire experts to manually label sponsored content in response to the potential damage caused by unreported sponsorship practice. In this regard, automatically detecting sponsored posts is essential for platforms to efficiently monitor user behaviors while reducing the workload of experts. However, research concerning the detection of sponsored content is still lagging behind [12–15]. One similar research is conducted in the context of fake reviews, where machine learning techniques are utilized to detect and classify fake reviews [16]. Unlike fake reviews, which mainly target at manipulating the reputation of products, sponsored posts are mainly applied for new product promotion, which aims to widely increase brand awareness, induce brand liking, and further strengthen customer purchase intention. Therefore, the sponsored posts are usually presented at longer length weaved with more complex linguistic expression and more obscure intentions.

In this study, we focus on the research question of how to detect whether a post is sponsored, and seek to delineate the underlying mechanism through the lens of advertising theory. To clearly illustrate the proposed research scenario and our key ideas, we hereby present several real-world examples from Xiaohongshu[1] in Fig. 1. In general, the posts on social media platforms can be classified into three categories: posts generated by sponsored SMIs, organic posts generated by users who do not receive any type of incentive, and posts generated by brand account. In this specific context, the goal is to recognize posts generated by sponsored SMIs, no matter whether the sponsorship is disclosed or not. To this end, we refer to relevant psychological theories and practical observations, and identify several useful clues including content overlap, image theme and sentiment shift. Based on these clues, we extract four distinctive features: textual, semantic, sentimental, and visual features through different approaches including Long Short-Term Memory (LSTM), sentiment analysis, text analysis, and image recognition. Moreover, given that different categories of products may prefer different types of features, we propose a unified deep learning detection framework with an adaptive weighting component.

To verify the effectiveness of our identified clues, extracted features, and proposed framework, we conduct extensive experiments on 26,823 social media posts collected from Xiaohongshu. The results demonstrate the performance of our proposed model compared with competitive alternatives and the value of each feature. The learned feature importance also enables deeper phenomena understanding of this task.

The rest of this paper is organized into four sections as follows. Section 2 reviews the previous literature on influencer marketing, sponsorship transparency and brand endorsement. Section 3 presents overall design of the research model and four types of features. Sections 4 discusses the experiment results, analyzes the contribution of each feature, and presents the visualization of the learned weight distribution. Section 5 concludes the paper by discussing the theoretical, methodology and managerial implications of our findings. Section 6 discusses the limitations of this paper and future research directions.

2. Related work

2.1. Influencer marketing

With the boom of social media, influencer marketing has swiftly flourished on Internet platforms and been prioritized by market practitioners to enhance business performance [17–19]. The core concept of influencer marketing lies in transferring the influencer-follower relationship into consumer positive attitude and purchase intention towards endorsed product.

As explained by social-capital theory, the strength of influencer-follower relationship can be measured as bridging social capital, which bridges the gap between follower's social worship and commercial purchase behavior [20]. In sponsorship practice, SMIs successfully cash in their accumulated social capital through collaborating with brands to endorse products [21]. In addition, compared with other types of endorsement formats (e.g., celebrity endorsement), the eWOM of SMIs is commonly being viewed as more authentic and credible by consumers. Prior research demonstrates that comparing with traditional advertisements, consumers are less likely to recognize persuasive intent in nontraditional advertising formats, such as social media [22]. Currently, research on influencer marketing is still in initial stage and can be largely categorized into two streams, as shown in our review of key studies in the domain (see Table 1). One stream mainly focuses on the key factors affecting the formation and growth of SMIs, such as influencer's social and physical attractiveness and perceived popularity [15,23]. Particularly, opinion leaders and para-social relationships have been identified as two focal components of influencer formation [24]. The other stream has employed surveys and experiments to explore factors affecting influencer marketing outcomes, including intention to interact [25], to recommend [26], and most importantly, to purchase based on influencers' recommendation [27,28].

2.2. SMI authenticity and sponsorship transparency

The para-social relationship between SMIs and followers lead followers to believe content that originates from SMIs is sincere, noncommercial-driven, and essentially different from marketer-initiated communication [22]. In most cases, to gain followers' trust and maintain their own credibility, SMIs choose to disclose the sponsorship behind its content to make it seem more natural and reliable. However, the exponential development of influencer marketing has led to the abuse use of sponsorship practices, as its commercial essence began to surface and has raised controversy and public critique as biased, unreliable, and even deceptive [31]. According to reactance theory, in general, people want to maintain their freedom of choice and refuse to be manipulated [32]. Therefore, it is assumed that people are more likely to resist a persuasion attempt when they perceive the persuasion as a threat to their freedom. Persuasion knowledge model further suggests that once consumers are aware of the persuasion attempt of a message, they will refine their attitudes and behaviors using their previously acquired knowledge, as either attending to the persuasion or resisting it. Furthermore, persuasion knowledge model also points out that such awareness can trigger criticisms about honesty, trustworthiness, and credibility from readers. Earlier research has found that, among adults, increased awareness of the persuasive nature of sponsored content in blogs can have negative consequences for the perceived quality and credibility of the blog and the influencer [33].

Following the growing concern that influencer marketing practice might arise criticism from consumers, researchers devoted special attention to the role of sponsorship disclosure in social media domain. Recently, researchers have shifted their attention to the necessity of sponsorship transparency and sought to confirm its role in attenuating consumer negative reactions [10,11,25]. Among these studies, one stream focuses on how consumers build their attitudes towards sponsorship transparency while the other explores the factors that moderates the negative effects of sponsorship disclosure. For example, Kim et al. [11] found that sponsorship disclosure increases suspicions about the reviewer's ulterior motives and decreases consumers' attitudes and purchase intentions when a review is positive. However, sponsorship disclosure does not hurt attitudes or purchase intentions when a review is negative.

2.3. Brand endorsed content

By offering direct/indirect monetary compensation, the brand-SMI sponsorship entails marketers to partially or fully determine the content of a sponsored post contractually. Brand encroachment refers to

[1] Xiaohongshu, also known as Little Red Book, is a social media and e-commerce platform in China.

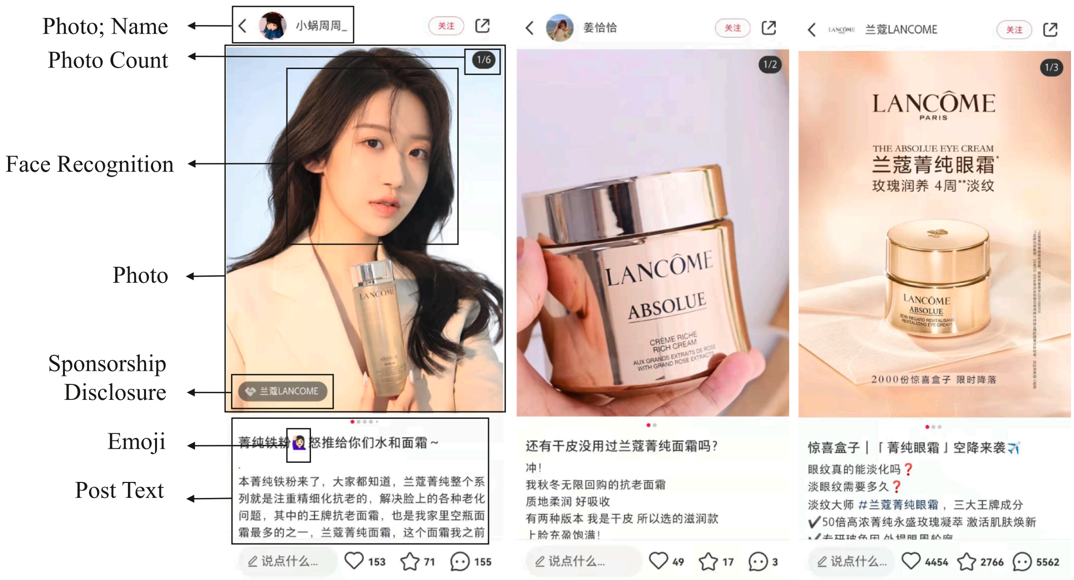


Fig. 1. An Example of Sponsored Post, Organic Post, and Brand-generated Post on Xiaohongshu App.

Table 1
Summary of recent studies on influencer marketing.

Category	Research objective	Factors	Ref.
SMIs attributes	Follower stickness	Identity similarity, identity distinctiveness, identity prestige	[13]
	Follower attitude	Attractiveness, likeability, similarity, closeness	[15]
IM outcomes	Follower engagement	Follower count, followee count, content volume, domains of interest	[14]
	Social media performance	Linguistic style, emotional contagion	[29]
	SMIs credibility	Expertise, trustworthiness, attractiveness	[23]
	Purchase intention	Opinion leadership, para-social relationship	[24]
	PSI, purchase intention	Photo types, gender, content generator types, branded content types	[27]
	Continuance follow intention	Influencer-product congruence, disclosure prominence	[25]
	Loyalty, product attitude, purchase intention	Source credibility, source attractiveness	[28]
	PSI, Luxury brand perceptions, purchase intention	Attitude homophily, physical attractiveness, social attractiveness	[30]
	Attitude toward recommendation, purchase intention	Sponsorship type, product type, brand awareness	[26]

embedding varying degrees of brand-related information into SMI content, including a redirection link, pictures featuring SMI with endorsed product, and minimum number of posts mentioning the products, etc. [34]. SMIs seamlessly blend such endorsed content with infectious expressions into their daily posts to disguise them as heartfelt recommendations [6]. As a consequence, such narrative technique blurs the line between sponsored posts and genuine praise that user-generated. Zhou et al. [6] found that bloggers utilize four narrative styles, including evaluation, embracing, endorsement, and explanation, to alter marketing messages in their sponsored eWOM to make the branded content more trustworthy, relevant, and useful to consumers. Previous studies on endorsed content detection are mainly conducted in the context of sponsored reviews, which are written by people with offered compensation to deliberately manipulate the valence of eWOM [35,36]. For example, Kim et al. [11] conducted text mining analysis on sponsored reviews and revealed that sponsored reviews provide more elaborate and evaluative content but are perceived as less helpful than organic reviews.

With a different purpose, sponsored posts are usually presented at a longer length and weaved with more product-specific elaboration to increase brand awareness and introduce new product. In this regard, marketers usually dictate specific requirements regarding the content (e.g., a minimum number of posts mentioning the products) to ensure consumers are aware of the product/brand. As a result, the detailed product information provided by the brand forms varying degrees of overlap among the content generated by sponsored SMIs and brand-generated posts. In contrast, such overlap rarely exists in organic posts since normal consumers lack motive and background knowledge to present detailed product information. Furthermore, the detailed product information also entails sponsored posts to cover a wider range of topics than organic posts.

As affect transfer theory states, an evaluation or feeling induced by an object can be transferred to another object [37]. In the context of influencer marketing, sponsored posts highlight SMIs' individual attractiveness and seek to transfer these types of attractiveness into the endorsed product [27,30]. As an effective modality to convey such attractiveness, visual content such as photos and videos have become a common format that SMIs frequently choose to achieve such process. In general, social media users posting product-related photos are motivated by providing visual clues for showing specific information from one's actual experience [27,38–40]. Besides, it also serves the purpose of verifying that the user has physically used the product and thus engendering trust [27,41]. In contrast, prior studies explored that SMIs post self-centric selfies and glamorous portraits to create substantial social influence for message recipients on social media, as manifested by increased followers' engagement [41,42]. Compared with sponsored reviews or organic posts, where the object of photos is usually the product or place (e.g., clothes, restaurant, and hotel), SMIs usually post self-promotional selfies or glamorous long-shot portraits with the endorsed product in it. The incorporation of portrait photos makes sponsored posts more persuasive to followers, but meanwhile, it also distinguishes them from organic posts or sponsored reviews.

3. Methodology

In this study, we propose a detection framework for sponsored posts on social media platforms by extracting the right clues from posts. However, experience goods and search goods on social media may not share the same set of clues[2]. For example, cosmetic products (experience goods) rely more on visual appearances of the influencers, compared with digital devices (search goods). In this research, we limit the scope to experience goods[3] because experience goods rely more on reviews and advertising [43], and most products in our research context (i.e., Xiaohongshu) are experience goods. We leave it as a future research direction to explore the search goods, as shown in Section 6.

The overall structure of the proposed model is outlined in Fig. 2. First, three types of posts: sponsored, organic, and brand-generated posts are collected using our android crawler. Based on psychological theories and practical observations, we identify several task-related cues which are helpful in detecting sponsored posts. Relying on the cues, we pre-process the posts and extract four types of features: textual, semantic, sentimental, and visual features through different approaches including Long Short-Term Memory (LSTM), sentiment analysis, text analysis, and image recognition. Finally, different features are integrated together to form the final feature representation of a post and then fed as input to the sponsored post detector. The details of the proposed detection framework will be discussed in the following sections.

3.1. Finding task-related clues

Given the advertising nature of sponsorship practice, the functions of sponsored posts can be characterized by two dimensions: informative and persuasive [44]. The informative function provides information to consumers about products, services, and prices while persuasive function cultivates purchase intention mainly through irrational means and peripheral clues. As the main component of a social media post, textual content takes the core informative role and deeply affects the reasoned decision-making choices of consumers. The psychology literature has analyzed that the comprehension of textual information can be processed in both micro-level (individual word) and macro-level (meaning of the document) [45]. To leverage the advantage of deep learning, this study utilized word embedding to learn word vector representations as well as the semantic relations between them.

In addition, as a distinctive feature of sponsored posts, the encroachment of brand-related content plays a key role in informative advertising. In practice, SMIs are contractually required to embed product-related information (e.g., material, ingredients) in their posts when cooperating with brands. As a consequence, the encroachment leads to varying degrees of overlap among the content generated by sponsored SMIs and brand counts. In contrast, such overlaps rarely exist in organic posts since normal consumers lack motive and background knowledge to present detailed product information. Therefore, we argue that it is necessary to incorporate additional information source: brand-generated posts for the identification of sponsored posts.

Besides informative function, sponsored posts also exhibit several clues that strongly indicate their persuasive intention. According to Aristotle's theory about the rhetorical forms of persuasive appeal [46], sentiment signifies the emotional appeals of communicators, and make readers more emotionally involved in the process of communication [6]. The sentiment of sponsored posts exhibits a distinctive pattern from other textual content (e.g., sponsored reviews), where the sentiment switches from extreme negative (previous bad experience) to extreme positive (product advantage) within a post. The sentiment shift could be utilized to serve the purposes of elicitation of sympathy and enhancing trustworthiness as well as contrasting the efficacy of endorsed product. Meanwhile, the sentiment shift greatly increases the sentiment complexity of sponsored posts, as some sponsored posts express almost no negative opinions while others contain even more negativity than organic posts. Therefore, instead of using document-level sentiment analysis [47,48], we incorporated aspect-level sentiment analysis in our study to further distinguish the sentiment orientation as experience-specific and product-specific sentiments.

Lastly, visual attractiveness plays a critical role for SMIs to appeal to followers, cultivate their purchase intention and persuade them to make imitative behaviors. Prior studies have suggested that SMIs constantly post self-centric selfies and portraits in order to maintain followers' engagement [38,40]. As affect transfer theory states, an evaluation or feeling induced by an object can be transferred to another object. In the case of this study, sponsored SMIs often post images of them holding or using endorsed products, indicating that their physical attractiveness is connected or partially attributed to the products. In contrast, while regular users also attach images to their content, these images mainly serve the purpose of informing and validating their argument. As a consequence, these photos are mainly product-related and no selfies are included [38,40,49]. In this study, we characterize the difference between two types of images by two statistical features: sponsored SMIs are more likely to post more images and the images are more self-centric. Specifically, the self-centric feature is captured by leveraging a deep learning face detection model.

3.2. Feature extraction

3.2.1. Textual feature extraction

To extract textual features, we employ recurrent neural networks (RNNs) with Long Short-Term Memory (LSTM) units to learn the representation of post text. RNNs are a type of feed-forward neural networks that can effectively model variable-length sequential information through the intermediate layers. However, RNNs suffer from problems of vanishing and exploding gradients, which leads to the model being inefficient in learning words dependencies with long-distance temporal. To overcome the problem of RNNs, LSTM extends the basic RNN with memory units and efficient gating mechanisms to store information over long time periods. An LSTM unit consists of a memory cell which keeps track of the dependencies between the word embedding and three gates: forget gate, input gate, and output gate. The forget gate determines what information to forget, the input gate decides what new information to update and the output gate combines the information retained in the previous cell with the new candidate information to form the final output. At each step in the citing sentence sequence, the LSTM takes both the last hidden state and the word embedding as the current input to compute the cell state vector C_t and the current hidden state with vector h_t, then creates a set of candidate values $\widetilde{C}_t$ to be added to the new cell state. The weight matrices W and bias vectors b, which need to be learned during training, determine how the gates operate. The computational learning steps taken in the LSTM unit are calculated as follows:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

$$\widetilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

[2] Experience goods are commodities where price, quality or some other attribute remains unknown until purchase, e.g., cosmetic, and jewelry), whereas search goods are commodities that have attributes that the buyer can evaluate before purchasing, e.g., automobiles and digital devices [43].

[3] Note that, the products are not neatly bifurcated into experience goods and search goods. Although the product categories we used in this study belong to experience goods, they have the attributes of search goods with different levels.

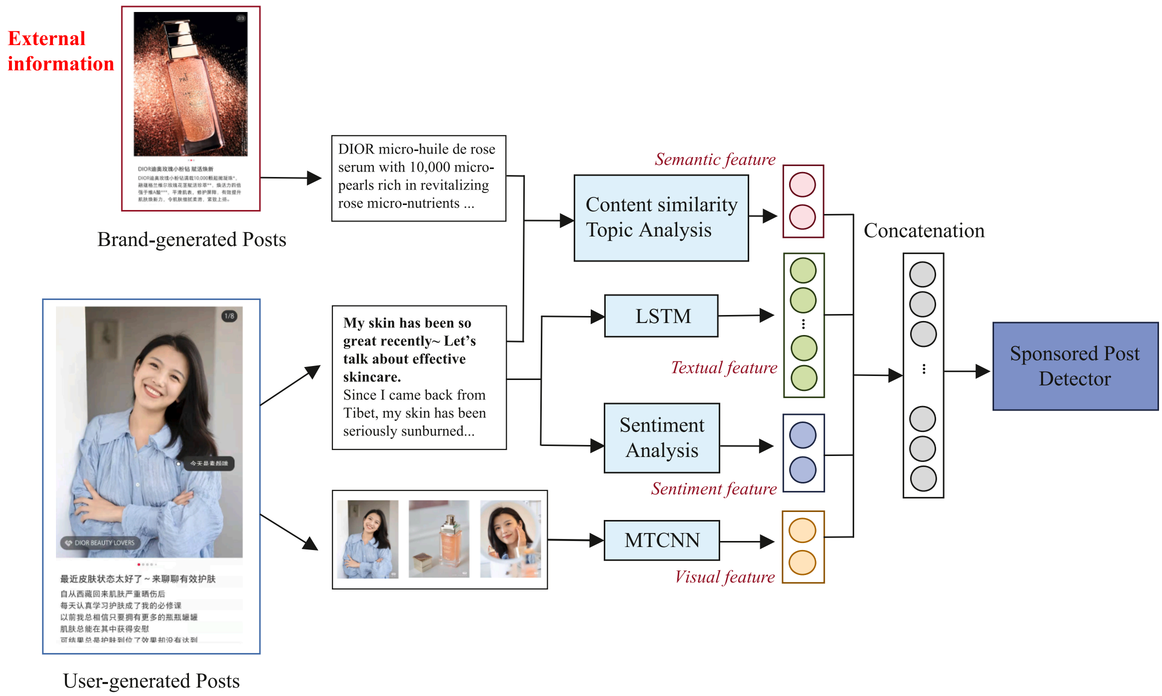


Fig. 2. The Proposed Model Framework.

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

where W_f, W_i, W_C, and W_o represent the weight matrices associated with each unit, and b_f, b_i, b_C, and b_o are the bias terms related to each unit.

Specifically, we employ stacked bi-directional LSTM units to extract textual features. The final hidden state of LSTM is obtained by concatenating the forward and backward states. Finally, we pass the LSTM output through a fully connected layer to get the textual features.

$$R_T = \phi(W_{tf} R_{lstm}) \tag{7}$$

where, R_{lstm} is the output of the LSTM, W_{tf} is the weight matrix of the fully connected layer and ϕ is the activation function used.

3.2.2. Content similarity

Previous studies on online reviews have found that semantic features play an important role in predicting review helpfulness, product sales and fake review detection [50,51]. However, it is worth noting that few research has yet investigated the difference between sponsored post and organic posts on a semantic level. As analyzed in Section 2.3, the brand encroachment leads to varying degrees of overlap between the content of sponsored and brand-generated posts. To measure the semantic similarity between sponsored posts and brand-generated posts, we utilize Doc2bow method in Gensim [52] to convert documents into Bag of Words representation. Cosine similarity is the calculation of the similarity between two n-dimensional vectors by looking for a cosine value from the angle between the two and is often used to compare documents in text mining. The content similarity between post P_i and its corresponding brand-generated post B_i is defined as follows:

$$\mathit{Content\ similarity}(P_i, B_i) = \frac{\overline{P_i} \bullet \overline{B_i}}{\|\overline{P_i}\| \ \|\overline{B_i}\|} \tag{8}$$

3.2.3. Topical analysis

Comparing with organic posts, sponsored posts are usually presented at a longer length and cover a wider range of topics to introduce product information in detail and foster customer purchase intention [6,53]. In this study, we adopted n-gram Latent Dirichlet Allocation (LDA) [54] to identify topic distribution in posts and utilized the perplexity scores to assess the optimal number of topics. LDA is a topic modelling approach that is widely used in the computer science and engineering fields to identify latent topic information from documents [50,55]. The core idea of LDA is that each document can be viewed as a representation of a probability distribution of topics, and each topic can be viewed as a representation of a probability distribution of many words. The result of topic distribution indicates that the topics of posts with wide topic coverage are evenly distributed while posts with a single topic always present a pattern, that is, the proportion of one topic is particularly high while the proportion of other topics is fairly low. Let $T_i = [t_{i1}, \ldots, t_{in}]$ be the topic distribution vector of post P_i that we calculated through LDA, the topic coverage of a post is defined as follows:

$$\mathit{Topic\ coverage}(P_i) = \sqrt{\frac{n-1}{\sum_{k=1}^{n}\left(t_{ik} - \overline{T_i}\right)^2}} \tag{9}$$

3.2.4. Aspect-level sentiment analysis

Sentiment analysis has been widely used to explore consumers' attitude or emotion towards products in all types of user-generated content (UGC) [56,57]. Sentiment analysis on UGC can be broadly categorized into four types: lexicon-based, statistics-based, rule-based, and machine learning approaches [58,59]. In this study, we adopted lexicon-based method to serve the goal of not only classifying sentiment orientation, but also calculating the sentiment intensity of each post. We apply the Chinese sentiment dictionary National Taiwan University Sentiment Dictionary (NTUSD) due to its higher accuracy than other lexicons in our pre-experiment.

Currently most sentiment analysis methods are conducted on document-level, which ex- tracts the sentiment of the whole given document, such as a review or a tweet. However, after analyzing the sentiment distribution pattern of sponsored and organic posts, we found that the document-level sentiment analysis only is not sufficient to cope with our task. Based on the result of document-level sentiment analysis, a substantial proportion of sponsored posts contain even more negative sentiment than organic posts, while others express rare negative orientation. Such distribution inconsistency makes sentimental feature difficult to be a distinguishing indicator, and even brings more noise to final detection model. Based on these observations, we found that the negative content is usually expressed in the beginning of a sponsored post. This phenomenon can be explained by a writing technique SMIs utilize, that is, deliberately emphasizing their previous bad experience to contrast the efficacy of the product they promote. Therefore, we chose aspect-level sentiment analysis in our study to distinguish the sentiment orientation as experience-specific sentiment and product-specific sentiment. We adopted the approach utilized in [60], which uses the probability value of a word in a topic generated over the topic-word distributions as the term weight of the word. For this model, the sentiment score of topic A is calculated as

$$Sentiment\ score(A) = \sum_i Prob(word_i|Topic_A) \times SO(words) \tag{10}$$

A topic sentiment score is calculated by multiplying the SO value by the probability of words and summing the products in a given topic, and then divided by the number of total words in a post. Further, the same topical words could be assigned in multiple topics with different probability values.

3.2.5. Visual analytics

As analyzed in Section 2.3, SMIs usually post self-promotional selfies or glamorous portraits with endorsed product in it to achieve persuasion implicitly. In this study, based on existing studies and observations, as shown in Table 2, two phenomenal features are captured: total photo count and portrait photo count. We transferred visual information into statistic features instead of multidimensional embedding mainly because the features of sponsored photo are not easy to align with either organic photos or textual content. Therefore, to obtain the portrait photo count, we choose a deep learning approach named multi-task cascaded CNNs (MTCNN) based framework [61] for the face detection task. Like the architecture depicted in Fig. 3, MTCNN adopts a cascaded structure with three stages of carefully designed deep convolutional networks that predict face and landmark location in a coarse-to-fine manner.

3.3. Sponsored post detector

In this subsection, we introduce the sponsored post detector. s shown in Table 2, we have obtained a representation P_T for text and numerical representation for sentiment P_{ST}, semantic P_{SM} and visual P_V. Given that the four types of features are of different contribution to final classification results, a weight vector of different features is denoted as $[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$ and to be learned in the model training phase. These features are then concatenated to form the representation for a given post as follows:

$$P_m = [\alpha_1 \cdot P_T, \alpha_2 \cdot P_{ST}, \alpha_3 \cdot P_{SM}, \alpha_4 \cdot P_V]. \tag{11}$$

Table 2
Four Types of Features Extracted in This Study.

Features	Description	Ref.
Textual-based features	Word embedding of post text	[62]
Semantic-based features		
Content similarity	Similarity between a post and corresponding brand post	[63]
Topic coverage	Topic coverage of a post	[64]
Sentiment-based features		
Experience-specific sentiment	Experience-aspect sentiment score	[65]
Product-specific sentiment	Product-aspect sentiment score	[66,67]
Visual-based features		
Photo count	Total no. of photos in a post	[27,38–40, 49]
Portrait photo count	Total no. of portrait photos in a post	[5,27,41,42, 68]

Then the concatenated features are fed into a softmax layer to calculate the probability of this review being sponsored. We denote the sponsored post detector as $G_d(P_m, \theta_d)$, where θ_d represents all the parameters included. The output of the sponsored post detector for a post P_m is the probability $\hat{y}_m$ of this post being a sponsored post.

$$\hat{y}_m = G_d(P_m, \theta_d) \tag{12}$$

In order to constrain the values of $\hat{y}_m$ within range of [0,1], we use the sigmoid logistic function. Finally, we employ the cross-entropy to calculate the loss of post P_m as follows.

$$\mathscr{L}(P_m) = -\mathbb{E}_{(m,y)\sim(M,Y)}[y\log(\hat{y}_m) + (1-y)\log(1-\hat{y}_m)] \tag{13}$$

where M represents the set of multimedia posts and Y represents the set of ground truth labels. We minimize the detection loss function $L_d(\theta_f, \theta_d)$ by seeking the optimal parameters $\hat{\theta}_f$ and $\hat{\theta}_d$, and this process can be represented as:

$$(\hat{\theta}_f, \hat{\theta}_d) = \underset{\theta_f,\theta_d}{\operatorname{argmin}} G_d(\theta_f, \theta_d) \tag{14}$$

4. Results and analyses

4.1. Dataset

The data were collected on Xiaohongshu, a rapid growing social commerce platform in China, which has more than 300 million users by the end of 2019. The majority of users on the platform are urban white-collar women with strong consumption demand, who can generate more than 3 billion posts every day around topics including cosmetics, fashion, baby_care, travel and fitness, etc. As a consequence, the enormous user base and their potential purchasing power makes Xiaohongshu an ideal channel for brands to conduct influencer marketing and cultivate brand awareness. Nevertheless, Xiaohongshu is one of the first platforms in China to implement sponsorship disclosure mechanism that requires all users to report if their post is sponsored.

To obtain research data, an Android crawler in Python is developed to collect data from Xiaohongshu app from February 1 to July 31, 2021. We continuously collected 26,823 posts related to eleven products under three categories: cosmetics, jewelry, and baby_care. The posts were collected by clicking product-related hashtags and then we crawled all posts presented on the topic page, which are published by influencers, regular users as well as brand official accounts. In addition, Xiaohongshu has adopted both machine and manual scrutiny mechanism to all user-generated posts and conducted an additional scrutiny process for posts with hashtags. All detected sponsored posts are being labelled in metadata information, which is crawled and leveraged as the label system of this study. The strict scrutiny process of the platform ensures the accuracy of collected ground truth labels. The posts crawled by our program show clear unequal statistical values between sponsored and organic posts. The descriptive statistics is presented in Table 3, including mean value, standard deviation and maximum values of two types of posts. Furthermore, to confirm the distinctiveness of types of features, we performed an independent-sample t test on the feature value of two types of posts.

4.2. Experimental design

The primary objectives of this study are to delineate the unique features of sponsored posts and build a multimodal[4] deep learning

[4] Although the major intended contribution is not designing a novel multimodal learning framework, we use the term multimodal to reflect that our research utilizes multiple modalities of the posts as data sources. See Section 5.2 for the details of contribution.

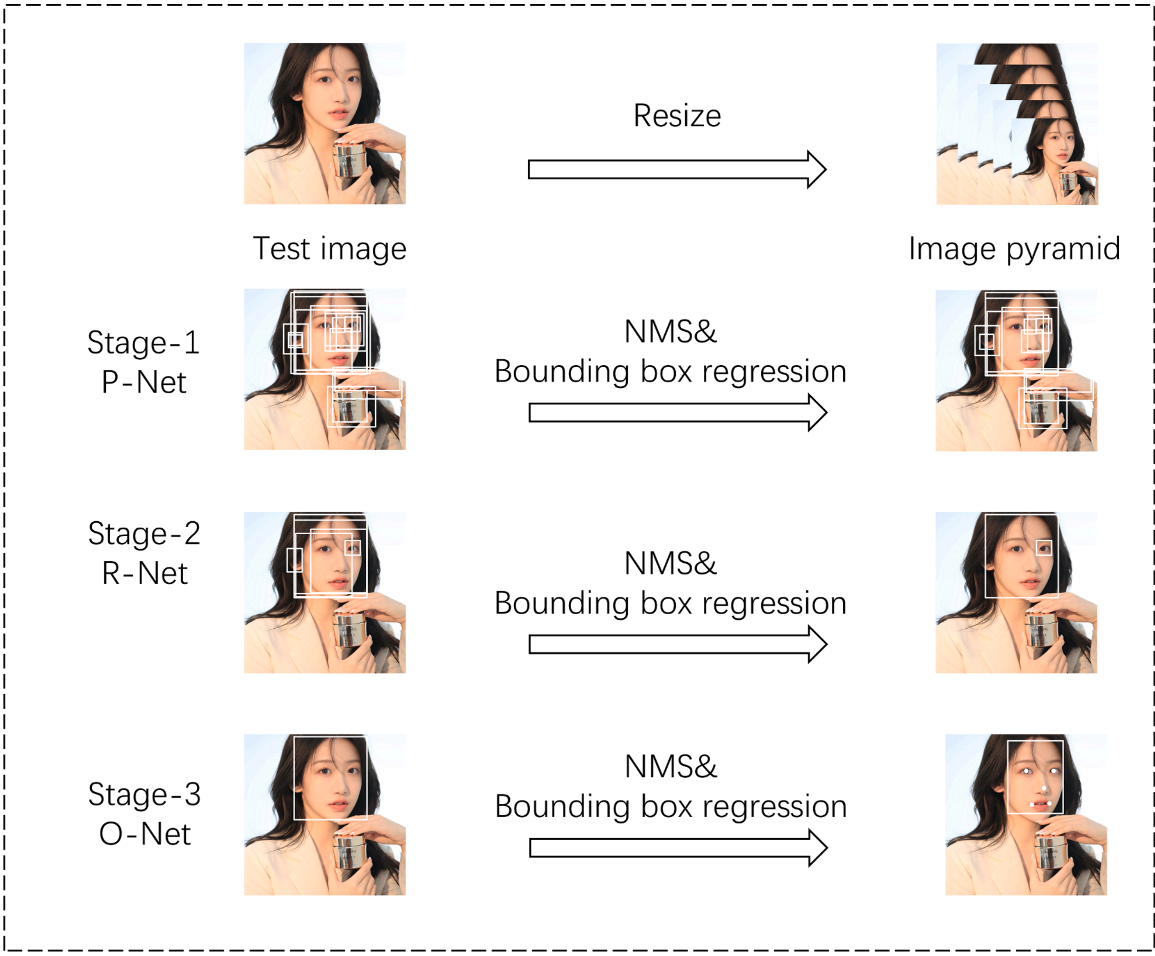


Fig. 3. MTCNN Algorithm Framework.

Table 3
Descriptive Statistics of Post Features and Results of Independent-samples *T* Test.

Feature Types	Feature	Original			Sponsored			T statistic
		Mean	Std.	Max	Mean	Std.	Max	
Semantic	Content similarity	0.14	0.19	0.56	0.32	0.21	0.98	26.17***
	Attribute coverage	0.01	0.19	0.20	0.12	0.18	1.00	5.02***
Sentimental	Experience-specific sentiment	0.08	0.06	0.34	0.05	0.04	0.17	− 1.74**
	Product-specific sentiment	0.14	0.13	0.58	0.24	0.12	0.53	7.38***
Visual	No. of photos	3.30	2.39	9	7.72	1.61	9	30.88***
	No. of portrait photos	0.55	1.18	8	3.00	1.94	8	24.40***

framework (SSV-LSTM) for efficient feature integration. Correspondingly, our evaluation is twofold: effectiveness of our proposed framework (model-level) and the value of each feature (feature-level).

First, in the proposed model, we leverage both text and image information to detect sponsored posts. Since textual content constructs the key component of social media posts, it can also be solely used to detection sponsored posts. Thus, to evaluate the effectiveness of the proposed approach, we compare it with two categories of baseline models: single modal models and multimodal models:

- TextCNN [69]: TextCNN is a well-known CNN-based classification approach based on textual input.
- LSTM [70]: LSTM is a special type of RNN comprising of various gates and obtains superior performance for text classification.
- SVM [71]: We leverage word2vec embedding as textual features and concatenate three types of features as additional input into Linear SVM.
- EANN [72]: The Event Adversarial Neural Networks (EANN) model consists of three components: multimodal feature extractor, fake news detector and event discriminator. In order to use it as a baseline, we remove the event discriminator module and take the first image as input of visual feature. Given that the task of sponsored post detection is less studied, to the best of our knowledge, EANN is the one of the most representative and applicable models for benchmarking.

Second, since feature engineering constructs the key component of the proposed approach, we also compared the performance of classification models with different feature input. Specifically, for sentiment features, we compare the classification performance based on document-level sentiment (Doc_Sen), aspect-level sentiment extracted in this study (Asp_Sen), and their combination with textual features. For visual features, given that the multi-image issue of sponsored posts is challenging and lacks targeted solutions, we present the image classification results of a state-of-the-art image classification model: Resnet. The experiments of single feature are conducted on SVM framework and experiments of feature combinations are conducted on the proposed framework separately.

The performance of classifier models was evaluated by four metrics including average accuracy of three datasets and precision, recall, and F1 measure of each dataset. The evaluation of model performance was based on 10-fold cross validations. Specifically, we ran each detection model 10 times, and reported the average performance of the models.

4.3. Results

We evaluate the algorithm performance by distinguishing sponsored posts in terms of accuracy, precision, recall, and F1 score. A comprehensive performance evaluation of five classification models is provided in Table 4. As shown in Table 4, SSV-LSTM achieved the best performance on three datasets in terms of accuracy, precision, and F1 score. Specifically, the proposed method outperforms single modal model LSTM by 5.7% in terms of F1 score since SSV-LSTM can leverage the benefits of multimodal feature learning. Additionally, the proposed method also performs better than multimodal model EANN by 4.9% in terms of F1 score, which may be due to the fact that features in SSV-LSTM are more representative than automatically extracted ones in EANN. Moreover, paired-sample *t*-tests on performances between baseline methods and the proposed approach demonstrated that SSV-LSTM significantly outperformed the baseline models ($p < 0.05$).

Table 5 presents the performance comparisons among classification models based on different feature inputs. The results show that, whether as a single input or combined with textual features, the sentiment features and visual features extracted in this study are more effective than the other two feature extraction practices. Specifically, Asp_Sen feature outperformed Doc_Sen feature by 6.9% in terms of F1 score, which validates that aspect-level sentiment is more indicative than overall sentiment in sponsored posts context. Moreover, the unsatisfactory performance of Resnet could be attributed that multi-image issue of sponsored posts is fairly complex, where the images of posts are of unfixed number (e.g., 1 to 9 in this study) and different objects (e.g., product, influencers, receipt). Regarding this, the results demonstrate that the statistic features extracted in this study provide an effective solution for this issue.

Moreover, we conducted paired-sample *t*-tests on performances between the proposed SSV-LSTM model and the variations of SSV-LSTM (Sentiment-LSTM, Semantic-LSTM and Visual-LSTM). The results of *t*-tests demonstrate that SSV-LSTM significantly performed better than its variations except for the comparison to Semantic-LSTM on cosmetics dataset.

4.4. Discussion of the learned weights

To check the generalizability of the proposed SSV-LSTM approach and gain more insights into the contribution of features on products of diverse characteristics, we conducted repetitive experiments and present the visualization of learned weights in Fig. 4. As shown in Fig. 4, the weight of each feature was initially set to 0.25 and converged to a stable weight distribution in repetitive experiments. As expected, the textual feature plays the most important role across all product categories since text is the main component of a post. One noticeable difference is that besides textual feature, semantic features contribute most in cosmetics and baby_care datasets while visual features contribute most in jewelry dataset. Specifically, when products focus less on appearance but more on functionality (baby_care > cosmetics > jewelry, in terms of functionality), the weight of visual feature decreases while the weight of semantic feature increases. This is consistent with the fact that informative advertising is more suitable for functional products while persuasive advertising is more suitable for products focusing on visual attributes. Overall, the visualization of the learned weight distribution demonstrates that our approach can adaptively learn the weights of different features and possess the generalizability to fit in different category scenarios.

4.5. Case study

To illustrate how each type of feature contributes in sponsored post detection, we provide a case study by performing ablation experiments among the proposed SSV-LSTM model and SSV-LSTM model removing each type of feature. Fig. 5 presents three example sponsored posts, where Figs. 4a, 4b, and 5c correspond to posts that are successfully detected by SSV-LSTM but missed by removing semantic, sentiment, and visual features, respectively. For simplicity and clarity, only representative part of post text and the first image are presented in these examples.

Fig. 5a represents a type of sponsored post whose advertising attempt is linguistically well-disguised and thus hard to discriminate by considering simple textual feature. However, as the red color text in Fig. 5a shows, these sponsored posts are still inevitably constrained by the encroachment requirements (e.g, product-related information insertion). In comparison, as discussed in Section 2.3, normal consumers lack of motive and background knowledge to introduce detailed product information. Hence, comparing the text of sponsored posts with brand-generated posts could effectively capture such advertising nature. In the posts detected by incorporating sentiment features, as shown in Fig. 5b, the SMIs tend to start their posts by emphasizing their negative experience or condition, and then elaborate the advantages of endorsed products. The combination of two polarized sentiment could attenuate the indicative effect of sentiment feature, and in this study we decompose the overall sentiment into two aspects: experience-specific and product-specific. Fig. 5c represents the sponsored posts that are additionally detected by incorporating visual features. The text of these posts shows little evidence of advertising attempt, but instead these posts focus more on creating attractive images so as to achieve influencer marketing.

5. Conclusions and discussions

The boost of online social media and the emergency of SMIs has profoundly impacted the form of online marketing but also bred the problem of trust crisis. Guided by psychological theories and real-world observations, this research characterizes the advertising nature of

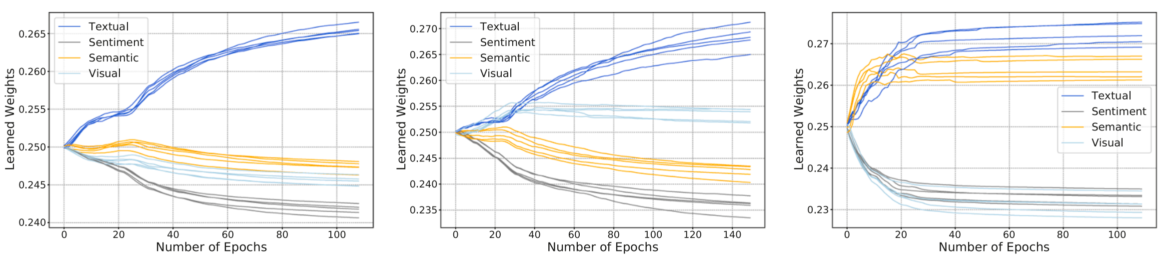
Table 4
Comparative Performance of Different Methods on Three Datasets.

Method	Accuracy	Cosmetics			Jewelry			Baby_care		
		Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
TextCNN	0.845	0.928***	0.838***	0.881***	0.875***	0.597***	0.710***	0.867***	0.578***	0.693***
LSTM	0.852	0.869**	0.981**	0.922**	0.918***	0.620***	0.740***	0.756**	0.656***	0.702***
SVM	0.848	0.872***	0.944***	0.907***	0.821***	0.691***	0.751***	0.759**	0.667***	0.710***
EANN	0.839	0.880**	0.981***	0.928**	0.841**	0.710*	0.750*	0.750*	0.667***	0.706*
SSV-LSTM	0.876	0.933	0.972	0.952	0.900	0.752	0.820	0.744	0.711	0.727

*, ** and *** denote significance at the.05,.01 and.001 levels, respectively.

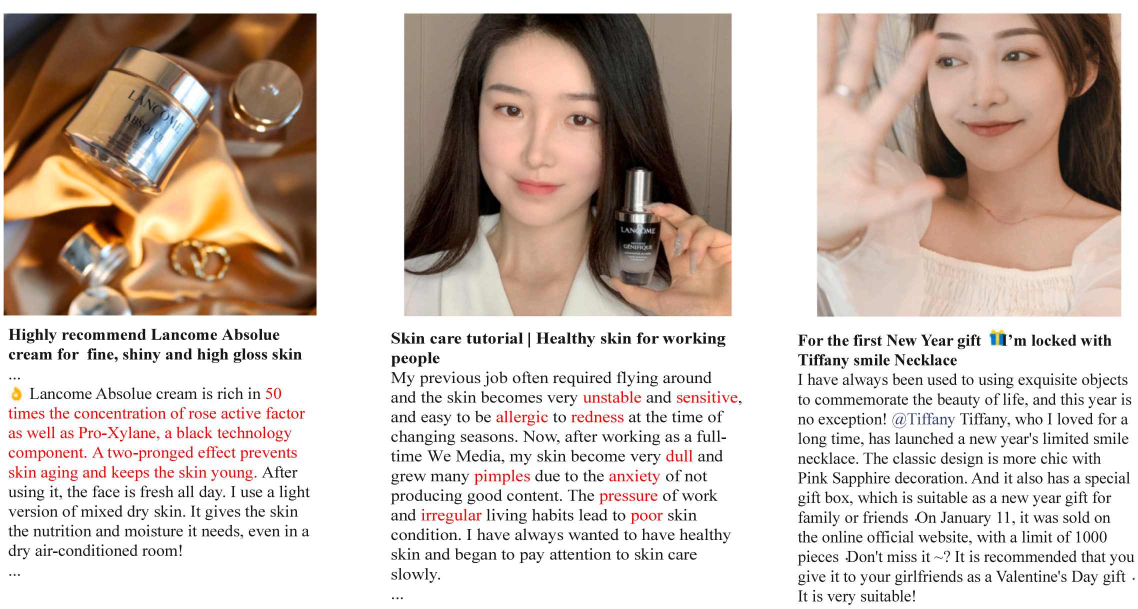
Table 5
Comparative Performance of Classification Models Based on Different Features.

View	Accuracy	Cosmetics			Jewelry			Baby_care		
		Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Doc_Sen	0.775	0.698	0.909	0.789	0.773	0.612	0.665	0.548	0.680	0.607
Asp_Sen	0.806	0.815	0.880	0.846	0.857	0.600	0.706	0.640	0.667	0.653
Textual+Doc_Sen	0.820	0.876	0.949	0.911	0.693	0.759	0.725	0.595	0.815	0.688
Textual+Asp_Sen	0.854	0.889	0.963	0.924	0.752	0.759	0.756	0.814	0.633	0.713
Semantic	0.826	0.840	0.826	0.880	0.752	0.734	0.724	0.611	0.846	0.710
Textual+Semantic	0.867	0.932	0.949	0.940	0.862	0.722	0.786	0.808	0.656	0.724
Resnet	0.735	0.742	0.605	0.667	0.522	0.316	0.393	0.500	0.455	0.476
Visual	0.811	0.848	0.882	0.862	0.884	0.648	0.746	0.654	0.599	0.622
Textual+Visual	0.875	0.884	0.986	0.932	0.838	0.766	0.800	0.885	0.600	0.715



(a) Cosmetics (b) Jewelry (c) Baby_care

Fig. 4. Learned Weight Distribution for Each Category.



(a) Incorporate sentiment feature (b) Incorporate semantic feature (c) Incorporate visual feature

Fig. 5. Examples of Sponsored Posts Detected by SSV-LSTM but Missed by Removing One Type of Feature.

sponsored posts and identifies several relevant task-related clues. Based on the clues, we extract four types of features: textual, semantic, sentiment, and visual. A unified deep learning framework is proposed to adaptively learn the weights of different features in integration module, which enables the proposed framework the generalizability to fit in different category scenarios. Our experiments were conducted on 26,823 online social media posts collected from Xiaohongshu, a rapid growing social commerce platform in China. The evaluation results demonstrate that incorporating these four features yielded a significant improvement in sponsored post detection performance. The findings of this research provide several novel research contributions and practical implications for sponsored post detection.

5.1. Theoretical contribution

Despite the raised concern about the dark side of sponsorship practice [10,11], there is still a paucity of applicative research on the detection of sponsored content. In this study, we focus on a novel real-world problem—sponsored post detection, which possesses its unique challenges and features compared with other relevant applications. Viewed through the lens of advertising theory, our study reveals the advertising nature of sponsorship posts and confirms two hidden intentions: informative and persuasive advertising [44]. From the perspective of applications, the encroachment of brand-related content is found to be a unique feature that derives from the purpose of informative advertising. Additionally, this study demonstrates that incorporating brand-generated posts is efficient in capturing brand encroachment despite SMIs' effort in linguistically disguising it. By doing so, this study delivers applicative contribution on the algorithm design for sponsored post detection. Moreover, the automatic weighting module of the proposed framework also enables us to understand the relative importance of clues across different product categories.

Our research, in turn, contributes to and provides empirical evidence for the relevant psychological theories on which this research is based. Specifically, we draw upon rhetorical theory and affect transfer theory to understand how sponsored posts achieve persuasive advertising through sentiment and visual and routes. The sentiment pattern of sponsored is unique from other textual content (e.g., sponsored reviews) since it could switch from extreme negative to extreme positive within a post. The understanding of such pattern is rooted in the rhetorical persuasive appeal from Aristotle's theory, as sentiment shift may get readers more emotionally involved in the process of communication. Moreover, there is a growing interest in understanding the underlying mechanism of how to transfer followers' admiration for SMIs into purchase intention of endorsed product [4,27,30]. In this study, the number of portrait photos posted by SMIs is demonstrated to be an essential medium in underpinning this transfer mechanism. In this regard, this study provides an extension to and support of affect transfer theory.

5.2. Methodological contribution

This study also makes several methodological contributions in dealing with sponsored posts. First, while unreported sponsored content has bred the problem of trust crisis, there has been little research on how to build effective detection models for sponsored posts. This study is one of the first to propose a sponsored post detection framework that incorporates information from multiple data source and modalities, and combines state-of-the-art machine learning algorithms. Second, this study contributes to relative literature by suggesting that aspect-level sentiment analysis is more effective than document-level ones in dealing with sentiment of sponsored posts. Third, additional information source: brand-generated posts are incorporated in this study and demonstrated to be one of the most crucial information for enhancing the performance of sponsored post detection. This study reveals the endorsement nature of sponsored posts and thus provides methodology implications on which information should be leveraged in such tasks. Fourth, given that the unfixed number as well as diversified subjects of post photos, the extraction of visual features under sponsored posts context is quite different from that in other domains. In this regard, this study not only experimentally demonstrates the critical role of photos in sponsored post detection task, but also goes a step further by proposing an effective method of dealing with these photos. Last, this study proposes an automatic feature weighting component which fits our scenario and delivers valuable insights to enhance phenomena understanding.

5.3. Managerial implications

This study also delivers actionable insights for relevant stakeholders including platform managers and consumers. According to a recent survey, nearly 40% of consumers complain about seeing too much branded content on social media and 48% of them are starting to distrust influencers [73]. From a practical standpoint, this study suggests practical implications for platform managers who have struggled with insufficient regulation over the widespread sponsorship practice. Although manual checking has been introduced on some platforms as a corresponding strategy, it comes at the expense of human resources which is of high cost and constraints. Moreover, as the platform continues to grow, it is expected that the current strict scrutiny process will become more time-consuming. Accurate detection of sponsored posts can help the platform filter for posts that need further manual checking, hence saving lots of money for the platform. In addition, compared with machine security, manual checking is much more delayed, and the authors might have to wait for a long time before their posts pass the manual checking. Accurate detection of sponsored posts can help the platform design different scrutiny schemes for posts with different levels of confidence, finally improving the user experience and fostering the creation of a healthier social media ecosystem.

Additionally, this study also delivers implications for consumers by engaging critical thinking in the current era of ubiquitous online falsehood. As potential beneficiaries as well as victims of influencer culture, it is vital for consumers to make a nonintuitive judgment and avoid being entertained and deceived simultaneously. In this regard, the findings of this study provide practical clues for consumers to distinguish sponsored content in terms of textual, semantic, sentiment, and visual aspects. Particularly, when reading social media posts, consumers should be cautioned of the hidden advertising intention behind those carefully designed images and emotion-provoking words.

6. Limitations and further research

This research has several limitations and offers a number of directions for future research. In this research, we limit the scope to experience goods because experience goods rely more on reviews and advertising [43], and most products in our research context (i.e., Xiaohongshu) are experience goods. However, experience goods and search goods on social media may not share the same set of clues. Therefore, it is interesting to craft another set of psychologically motivated clues for search goods or conduct a comparative study between experience and search goods. In addition, the conceptualization and measurement of the semantic features still remains to be improved. Future researchers are encouraged to utilize deep learning approaches to learn better representations of the semantic pattern. Moreover, the inter-modality and intra-modality of textual and visual features can also potentially be improved through advanced fusion techniques of multimodal approaches. In this study we transformed the visual features into statistical characteristics instead of multidimensional embedding due to that we found the high dimensional image features is not as indicative as statistical features (e.g, portrait photo count). Recently, many multimodal or multiview representation methods utilize deep learning to learn the representative features and obtain superior performance through fusing different modalities [74]. In this regard, future research into the interaction and fusion of different modalities can potentially provide new insights into how SMIs conduct persuasion and expand social influence. Last, it is interesting to explore whether we could tap into the wisdom of crowds to further improve the performance [75]

CRediT authorship contribution statement

Rong-Ping Shen: Data curation, Writing – original draft, Writing – review & editing. **Dun Liu:** Conceptualization, Formal analysis, Methodology, Writing – review & editing. **Xuan Wei:** Conceptualization, Methodology, Writing – review & editing. **Mingyue Zhang:** Conceptualization, Methodology.

Acknowledgments

This work is supported by the National Science Foundation of China (Nos. 62276217, 61876157, 62136002, 72201167, 72272101), the Sichuan Science and Technology Program [No. 2022JDJQ0034], the Chongqing Key Laboratory Project of Computational Intelligence [No.

2020FF03], the Fundamental Research Funds for the Central Universities [No. 2682022ZTPY057], and the Shanghai Chenguang Program (21CGA13).

References

[1] H. Chen, W. Duan, W. Zhou, When products receive reviews across platforms: studying the platform concentration of electronic word-of-mouth, Inf. Manag. 58 (8) (2021) 103532, https://doi.org/10.1016/j.im.2021.103532.
[2] X. Cheng, Y. Gu, Y. Hua, X.R. Luo, The paradox of word-of-mouth in social commerce: exploring the juxtaposed impacts of source credibility and information quality on swom spreading, Inf. Manag. 58 (7) (2021) 103505, https://doi.org/10.1016/j.im.2021.103505.
[3] S.V. Jin, A. Muqaddam, E. Ryu, Instafamous and social media influencer marketing, Mark. Intell. Plan. 37 (5) (2019) 567–579, https://doi.org/10.1108/MIP-09-2018-0375.
[4] K. Sokolova, C. Perez, You follow fitness influencers on youtube. but do you actually exercise? How parasocial relationships, and watching fitness influencers, relate to intentions to exercise, J. Retail. Consum. Serv. 58 (2021) 102276, https://doi.org/10.1016/j.jretconser.2020.102276.
[5] E. Djafarova, C. Rushworth, Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users, Comput. Hum. Behav. 68 (2017) 1–7, https://doi.org/10.1016/j.chb.2016.11.009.
[6] S. Zhou, L. Barnes, H. McCormick, M.B. Cano, Social media influencers' narrative strategies to create eWOM: a theoretical contribution, Int. J. Inf. Manag. 59 (2021) 102293, https://doi.org/10.1016/j.ijinfomgt.2020.102293.
[7] W. Guo, K.J. Main, The vulnerability of defensiveness: the impact of persuasion attempts and processing motivations on trust, Mark. Lett. 23 (4) (2012) 959–971, https://doi.org/10.1007/s11002-012-9197-y.
[8] S. Kapitan, D.H. Silvera, From digital media influencers to celebrity endorsers: attributions drive endorser effectiveness, Mark. Lett. 27 (3) (2016) 553–567, https://doi.org/10.1007/s11002-015-9363-0.
[9] J. Singh, B. Crisafulli, M.T. Xue, et al., 'to trust or not to trust': the impact of social media influencers on the reputation of corporate brands in crisis, J. Bus. Res. 119 (2020) 464–480, https://doi.org/10.1016/j.jbusres.2020.03.039.
[10] Y. Hwang, S.-H. Jeong, "this is a sponsored blog post, but all opinions are my own": The effects of sponsorship disclosure on responses to sponsored blog posts, Comput. Hum. Behav. 62 (2016) 528–535, https://doi.org/10.1016/j.chb.2016.04.026.
[11] S.J. Kim, E. Maslowska, A. Tamaddoni, The paradox of (dis) trust in sponsorship disclosure: The characteristics and effects of sponsored online consumer reviews, Decis. Support Syst. 116 (2019) 114–124, https://doi.org/10.1016/j.dss.2018.10.014.
[12] A. Audrezet, G. de Kerviler, J.G. Moulard, Authenticity under threat: When social media influencers need to go beyond self-presentation, J. Bus. Res. 117 (2020) 557–569, https://doi.org/10.1016/j.jbusres.2018.07.008.
[13] L. Hu, Q. Min, S. Han, Z. Liu, Understanding followers' stickiness to digital influencers: the effect of psychological responses, Int. J. Inf. Manag. 54 (2020) 102–169.
[14] W. Tafesse, B.P. Wood, Followers' engagement with Instagram influencers: The role of influencers' content and engagement strategy, J. Retail. Consum. Serv. 58 (2021) 102303, https://doi.org/10.1016/j.jretconser.2020.102303.
[15] B.J. Taillon, S.M. Mueller, C.M. Kowalczyk, D.N. Jones, Understanding the relationships between social media influencers and their followers: the moderating role of closeness, J. Prod. Brand Manag. (2020), https://doi.org/10.1108/JPBM-03-2019-2292.
[16] G. Shan, L. Zhou, D. Zhang, From conflicts and confusion to doubts: examining review inconsistency for fake review detection, Decis. Support Syst. 144 (2021) 113513, https://doi.org/10.1016/j.dss.2021.113513.
[17] E. Uzunoğlu, S.M. Kip, Brand communication through digital influencers: leveraging blogger engagement, Int. J. Inf. Manag. 34 (5) (2014) 592–602, https://doi.org/10.1016/j.ijinfomgt.2014.04.007.
[18] D. Jiménez-Castillo, R. Sánchez-Fernández, The role of digital influencers in brand recommendation: examining their impact on engagement, expected value and purchase intention, Int. J. Inf. Manag. 49 (6) (2019) 366–376, https://doi.org/10.1016/j.ijinfomgt.2019.07.009.
[19] M. Zhang, X. Wei, G. Chen, Maximizing the influence in social networks via holistic probability maximization, Int. J. Intell. Syst. 33 (10) (2018) 2038–2057.
[20] S.-A.A. Jin, J. Phua, Following celebrities' tweets about brands: the impact of twitter-based electronic word-of-mouth on consumers' source credibility perception, buying intention, and social identification with celebrities, J. Advert. 43 (2) (2014) 181–195.
[21] M. De Veirman, V. Cauberghe, L. Hudders, Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude, Int. J. Advert. 36 (5) (2017) 798–828, https://doi.org/10.1080/02650487.2017.1348035.
[22] G. Van Noort, M.L. Antheunis, E.A. Van Reijmersdal, Social connections and the persuasiveness of viral campaigns in social network sites: persuasive intent as the underlying mechanism, J. Mark. Commun. 18 (1) (2012) 39–53, https://doi.org/10.1080/13527266.2011.620764.
[23] C. Pornpitakpan, The persuasiveness of source credibility: a critical review of five decades' evidence, J. Appl. Soc. Psychol. 34 (2) (2004) 243–281, https://doi.org/10.1111/j.1559-1816.2004.tb02547.x.
[24] S. Farivar, F. Wang, Y. Yuan, Opinion leadership vs. para-social relationship: key factors in influencer marketing, J. Retail. Consum. Serv. 59 (2021) 102371, https://doi.org/10.1016/j.jretconser.2020.102371.
[25] R. De Cicco, S. Iacobucci, S. Pagliaro, The effect of influencer–product fit on advertising recognition and the role of an enhanced disclosure in increasing sponsorship transparency, Int. J. Advert. 40 (5) (2021) 733–759, https://doi.org/10.1080/02650487.2020.1801198.
[26] L.C. Lu, W.P. Chang, H.H. Chang, Consumer attitudes toward blogger's sponsored recommendations and purchase intention: the effect of sponsorship type, product type, and brand awareness, Comput. Hum. Behav. 34 (2014) 258–266, https://doi.org/10.1016/j.chb.2014.02.007.
[27] S.V. Jin, E. Ryu, et al., I'll buy what she's #wearing: the roles of envy toward and parasocial interaction with influencers in Instagram celebrity-based brand endorsement and social commerce, J. Retail. Consum. Serv. 55 (102121) (2020), https://doi.org/10.1016/j.jretconser.2020.102121.
[28] D.Y. Kim, H.-Y. Kim, Trust me, trust me not: a nuanced view of influencer marketing on social media, J. Bus. Res. 134 (2021) 223–232, https://doi.org/10.1016/j.jbusres.2021.05.024.
[29] M.T. Lee, C. Theokary, A.G. Woodside, The superstar social media influencer: exploiting linguistic style and emotional contagion over content, J. Bus. Res. 132 (2021) 860–871, https://doi.org/10.1016/j.jbusres.2020.11.014.
[30] J.E. Lee, B. Watkins, Youtube vloggers' influence on consumer luxury brand perceptions and intentions, J. Bus. Res. 69 (12) (2016) 5753–5760. https://doi.org/10.1016/j.jbusres.2016.04.171.
[31] J. Kennett, S. Matthews, What's the buzz? undercover marketing and the corruption of friendship 1, J. Appl. Philos. 25 (1) (2008) 2–18, https://doi.org/10.1111/j.1468-5930.2008.00391.x.
[32] S.S. Brehm, J.W. Brehm. Psychological Reactance: A Theory of Freedom and Control, Academic Press, 1981.
[33] B.W. Wojdynski, H. Bang, K. Keib, B.N. Jefferson, D. Choi, J.L. Malson, Building a better native advertising disclosure, J. Interact. Advert. 17 (2) (2017) 150–161, https://doi.org/10.1080/15252019.2017.1370401.
[34] Y. Lu, M. Chau, P.Y. Chau, Are sponsored links effective? investigating the impact of trust in search engine advertising, ACM Trans. Manag. Inf. Syst. (TMIS) 7 (4) (2017) 1–33, https://doi.org/10.1145/3023365.
[35] S. Banerjee, A.Y. Chua, Calling out fake online reviews through robust epistemic belief, Inf. Manag. 58 (3) (2021) 103445, https://doi.org/10.1016/j.im.2021.103445.
[36] S. Lu, L. Qiu, K. Wang, The effects of the format of two-sided online reviews: a linguistic perspective, Inf. Manag. 58 (8) (2021) 103554, https://doi.org/10.1016/j.im.2021.103554.
[37] W.E. Baker, When can affective conditioning and mere exposure directly influence brand choice? J. Advert. 28 (4) (1999) 31–46, https://doi.org/10.1080/00913367.1999.10673594.
[38] Y.A. Argyris, Z. Wang, Y. Kim, Z. Yin, The effects of visual congruence on increasing consumers' brand engagement: an empirical investigation of influencer marketing on Instagram using deep-learning algorithms for automatic image classification, Comput. Hum. Behav. 112 (2020) 106443, https://doi.org/10.1016/j.chb.2020.106443.
[39] M. Kleemans, S. Daalmans, I. Carbaat, D. Anschütz, Picture perfect: the direct effect of manipulated Instagram photos on body image in adolescent girls, Media Psychol. 21 (1) (2018) 93–110, https://doi.org/10.1080/15213269.2016.1257392.
[40] E. Shin, J.E. Lee, What makes consumers purchase apparel products through social shopping services that social media fashion influencers have worn? J. Bus. Res. 132 (2021) 416–428, https://doi.org/10.1016/j.jbusres.2021.04.022.
[41] A. Rocamora, Personal fashion blogs: Screens and mirrors in digital self-portraits, Fash. Theory 15 (4) (2011) 407–424.
[42] B. Luvaas, What does a fashion influencer look like? Portraits of the instafamous, Fash. Style Pop. Cult. 4 (3) (2017) 341–364.
[43] P. Huang, N.H. Lurie, S. Mitra, Searching for experience on the web: an empirical examination of consumer behavior for search and experience goods, J. Mark. 73 (2) (2009) 55–69.
[44] P.C. Santilli, The informative and persuasive functions of advertising: a moral appraisal, J. Bus. Ethics 2 (1) (1983) 27–33.
[45] K.S. Goodman, Reading: a psycholinguistic guessing game, Lit. Res. Instr. 6 (4) (1967) 126–135.
[46] G.A. Aristotle, et al., Aristotle on rhetoric a theory of civic discourse(1991).
[47] R. Barbado, O. Araque, C.A. Iglesias, A framework for fake review detection in online consumer electronics retailers, Inf. Process. Manag. 56 (4) (2019) 1234–1244, https://doi.org/10.1016/j.ipm.2019.03.002.
[48] D. Zhang, L. Zhou, J.L. Kehoe, I.Y. Kilic, What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews, J. Manag. Inf. Syst. 33 (2) (2016) 456–481, https://doi.org/10.1080/07421222.2016.1205907.
[49] K. Lup, L. Trub, L. Rosenthal, Instagram# instasad? Exploring associations among Instagram use, depressive symptoms, negative social comparison, and strangers followed, Cyberpsychology Behav. Soc. Netw. 18 (5) (2015) 247–252, https://doi.org/10.1089/cyber.2014.0560.
[50] N.F. Ibrahim, X. Wang, A text analytics approach for online retailing service improvement: evidence from Twitter, Decis. Support Syst. 121 (2019) 37–50, https://doi.org/10.1016/j.dss.2019.03.002.
[51] S. Mitra, M. Jenamani, Obim: a computational model to estimate brand image from online consumer review, J. Bus. Res. 114 (2020) 213–226, https://doi.org/10.1016/j.jbusres.2020.04.003.

[52] R. Rehurek, P. Sojka, Software framework for topic modelling with large corpora, in: Proceedings of the LREC Workshop on New Challenges for NLP Frameworks, 2010.
[53] C. Lou, S. Yuan, Influencer marketing: how message value and credibility affect consumer trust of branded content on social media, J. Interact. Advert. 19 (1) (2019) 58–73, https://doi.org/10.1080/15252019.2018.1533501.
[54] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.
[55] D. Qiao, J. Zhang, Q. Wei, G. Chen, Finding competitive keywords from query logs to enhance search engine advertising, Inf. Manag. 54 (4) (2017) 531–543, https://doi.org/10.1016/j.im.2016.11.003.
[56] A. Levin. Influencer Marketing for Brands, Springer, 2020.
[57] J. Zhang, L. Wang, K. Wang, Identifying comparable entities from online question-answering contents, Inf. Manag. 58 (3) (2021) 103449, https://doi.org/10.1016/j.im.2021.103449.
[58] P. Turney, Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, 2002, pp. 417–424.
[59] T. Wilson, J. Wiebe, P. Hoffmann, Recognizing contextual polarity in phrase-level sentiment analysis. Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, 2005, pp. 347–354.
[60] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-based methods for sentiment analysis, Comput. Linguist. 37 (2) (2011) 267–307, https://doi.org/10.1162/COLI_a_00049.
[61] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, Joint face detection and alignment using multitask cascaded convolutional networks, IEEE Signal Process. Lett. 23 (10) (2016) 1499–1503.
[62] D. Shin, S. He, G.M. Lee, A.B. Whinston, S. Cetintas, K.-C. Lee, Enhancing social media analysis with visual data analytics: a deep learning approach, MIS Q. 44 (4) (2020) 1459–1492.
[63] S. De Jans, D. Van de Sompel, M. De Veirman, L. Hudders, # Sponsored! how the recognition of sponsoring on Instagram posts affects adolescents brand evaluations through source evaluations, Comput. Hum. Behav. 109 (2020) 106342, https://doi.org/10.1016/j.chb.2020.106342.
[64] H.J. Kim, Y.K. Jeong, Y. Kim, K.Y. Kang, M. Song, Topic-based content and sentiment analysis of Ebola virus on twitter and in the news, J. Inf. Sci. 42 (6) (2016) 763–781, https://doi.org/10.1177/0165551510000000.
[65] S. Mukherjee, P. Bhattacharyya, Feature specific sentiment analysis for product reviews. Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics, 2012, pp. 475–487, https://doi.org/10.1007/978-3-642-28604-9_39.
[66] R.K. Amplayo, S. Lee, M. Song, Incorporating product description to sentiment topic models for improved aspect-based sentiment analysis, Inf. Sci. 454 (2018) 200–215, https://doi.org/10.1016/j.ins.2018.04.079.
[67] W. Zhang, H. Xu, W. Wan, Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis, Expert Syst. Appl. 39 (11) (2012) 10283–10291, https://doi.org/10.1016/j.eswa.2011.09.135.
[68] A. McFarlane, E. Samsioe, # 50+ fashion Instagram influencers: cognitive age and aesthetic digital labours, J. Fash. Mark. Manag. Int. J. 24 (3) (2020) 399–413.
[69] Y. Zhang, B. Wallace, A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification, arXiv preprint arXiv: 1510.03820(2015). 10.48550/arXiv.1408.5882.
[70] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[71] J. Ma, W. Gao, Z. Wei, Y. Lu, K.-F. Wong, Detect rumors using time series of social context information on microblogging websites. Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, 2015, pp. 1751–1754.
[72] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao, Eann: event adversarial neural networks for multi-modal fake news detection. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 849–857.
[73] E. Dopson, 30+ influencer marketing statistics to have on your radar (2021), 2021, https://www.shopify.com/blog/influencer-marketing-statistics.
[74] Z. Zhang, X. Wei, X. Zheng, Q. Li, D.D. Zeng, Detecting product adoption intentions via multiview deep learning, INFORMS J. Comput. 34 (1) (2022) 541–556.
[75] X. Wei, Z. Zhang, M. Zhang, W. Chen, D.D. Zeng, Combining crowd and machine intelligence to detect false news on social media, Manag. Inf. Syst. Q. 46 (2) (2022) 977–1008.

Rongping Shen received the M.S. degree in software engineering from Southwest Petroleum University. She is currently pursuing her Ph.D. degree in Management Science and Engineering at the School of Economics and Management, Southwest Jiaotong University, Chengdu, China. Her research focuses on applying machine learning methods to business big data and contributing to the development of new methods of management problems. Her research results have been published in refereed journals and conferences, including Information Sciences, Expert Systems with Applications, IEEE International Conference on Big Knowledge (ICBK), etc.

Dun Liu received the B.S. and Ph.D. degrees from Southwest Jiaotong University, Chengdu, China in 2005 and 2011, respectively. He was a Post-Doctoral Researcher with the School of Economics and Management, Tsinghua University, Beijing, China from 2012 to 2014. He was also a visiting scholar at the University of Regina, Regina, SK, Canada in 2009 and Carnegie Mellon University, Pittsburgh, PA, USA in 2016. He is currently a Professor with the School of Economics and Management, Southwest Jiaotong University. Since 2005, he has published over 150 research papers in refereed journals and conferences, including IEEE Transactions on Fuzzy Systems, IEEE Transactions on Systems, Man, and Cybernetics: Systems, Fuzzy Sets and Systems, Information Sciences, International Journal of Approximate Reasoning, Applied Soft Computing, Computers & Industrial Engineering, Knowledge-based Systems, Expert Systems with Applications, etc. His current research interests include Uncertainty Modeling and Applications, Big Data Decision and Analysis, Three-way Decisions, Granular Computing, Information Management and Decision Support System.

Xuan Wei is an assistant professor in the Department of Information, Technology and Innovation, Antai College of Economics and Management, Shanghai Jiao Tong University. He received his B.S. degree from the Shanghai Jiao Tong University in 2014 and his Ph.D. degree in management information systems at the University of Arizona in 2020. His research interests include crowd intelligence and crowdsourcing, social media analytics, statistical machine learning, probabilistic modeling and interference, and deep learning. His work has been published in journals such as Nature Human Behaviour, MIS Quarterly, INFORMS Journal on Computing, Information & Management, etc.

Mingyue Zhang is an associate professor in the School of Business and Management, Shanghai International Studies University. She received her Ph.D. degree in management science and engineering from the School of Economics and Management, Tsinghua University in 2017. Her current research interests include online communities, incentive mechanisms, and social media analysis. Her work has been published in journals such as MIS Quarterly, Information & Management, Decision Sciences, Decision Support Systems, Information Sciences, ACM Transactions on Knowledge Discovery from Data, etc.