Seeking Micro-influencers for Brand Promotion

Tian Gan*
Shandong University
gantian@sdu.edu.cn

Shaokun Wang
Shandong University
wangskkk@163.com

Meng Liu
Shandong University
mengliu.sdu@gmail.com

Xuemeng Song
Shandong University
sxmustc@gmail.com

Yiyang Yao
Northwestern Polytechnical University
yao_yiyang@zj.sgcc.com.cn

Liqiang Nie
Shandong University
nieliqiang@gmail.com

ABSTRACT

What made you want to wear the clothes you are wearing? Where is the place you want to visit for your next-coming holiday? Why do you like the music you frequently listen to? If you are like most people, you probably made these decisions as a result of watching influencers on social media. Furthermore, influencer marketing is an opportunity for brands to take advantage of social media using a well-defined and well-designed social media marketing strategy. However, choosing the right influencers is not an easy task. With more people gaining an increasing number of followers in social media, finding the right influencer for an E-commerce company becomes paramount. In fact, most marketers cite it as a top challenge for their brands. To address the aforementioned issues, we proposed a data-driven micro-influencer ranking scheme to solve the essential question of finding out the right micro-influencer. Specifically, we represented brands and influencers by fusing their historical posts' visual and textual information. A novel $\mathcal{K}$-buckets sampling strategy with a modified listwise learning to rank model were proposed to learn a brand-micro-influncer scoring function. In addition, we developed a new Instagram brand micro-influencer dataset, consisting of 360 brands and 3,748 micro-influencers, which can benefit future researchers in this area. The extensive evaluations demonstrate the advantage of our proposed method compared with the state-of-the-art methods.

KEYWORDS

Multimodal, Influencer Marketing, Learning to Rank

ACM Reference Format:
Tian Gan, Shaokun Wang, Meng Liu, Xuemeng Song, Yiyang Yao, and Liqiang Nie. 2019. Seeking Micro-influencers for Brand Promotion. In *Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3343031.3351080

*Tian Gan is the corresponding author.

MM '19, October 21–25, 2019, Nice, France

ACM ISBN 978-1-4503-6889-6/19/10…$15.00
https://doi.org/10.1145/3343031.3351080

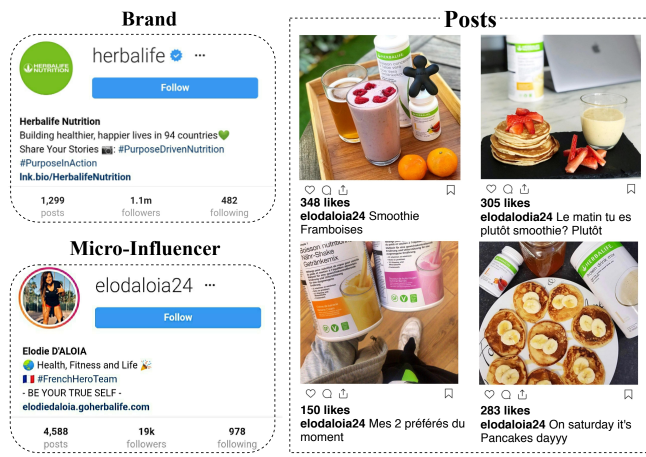

Figure 1: Example of a micro-influencer's advertizing posts.

1 INTRODUCTION

Social media has only become the mainstream for roughly over a decade, however, it has fundamentally impacted the way people conduct business. The ubiquity of social media allows businesses to easily access massive online social networks and interact with users on these networks [10, 24, 38]. According to a survey that, 83% of American adults use social media, with 48% interacting with brands and businesses on at least one social media site [4, 7, 28]. Obviously, this is a great opportunity for brands to market their products effectively in a short period of time.

However, the growing amount of useless content, such as unrelated or unreal advertisements, makes social media users more and more reluctant towards perceiving online advertisement. To mitigate such customer skepticism, marketers often leverage personalized content delivery channels, where influencer marketing clearly dominates over other marketing strategies. At the same time, consumers are more likely to trust brands that advertise via influencer channels rather than those who has adopted conventional marketing strategies [8]. Therefore, influencer marketing becomes a must and an essential element in advertising.

Prominent social media influencers are often celebrities including actors, writers, politicians, or athletes. In addition to celebrity influencers, there are a growing number of "non-celebrity" influencers on social media platforms [5], namely the micro-influencers. They are often referred to social bloggers with expertise in a particular product category and enjoying a considerable amount of followers on social media. However, most marketers tend to select the micro-influencers to be any YouTuber, Instagrammer, Snap chatter, or blogger with a relatively small (less than 100,000) follower-base [4, 28]. Since these micro-influencers are considered to be more cost effective, generate higher social media engagement, and have certain knowledge in the product domain and hence are more trustworthy compared to celebrity influencers [6, 17, 24]. For the reasons above, micro-influencers are treated as the best fit for the influencer marketing, and Figure 1 shows an example of a micro-influencer posting an advertisement about a brand.

In recent years, micro-influencer marketing has become an important element of social media marketing [1, 9, 11, 26]. However, relevant micro-influencers are difficult to find. With more and more people creating and expanding followers in social media, finding the right influencer for an E-Commerce company becomes paramount. That is why 75% of the brands feel that identifying the right micro influencer is the biggest challenge [35]. Also, the lack of open data for social influencer marketing analysis is also a barrier for research in this area.

To resolve the issues mentioned above, we proposed a multimodal micro-influencer ranking method, which leverages open data in social media to learn the relevance between brands and micro-influencers. In particular, we designed a social account history pooling mechanism that used posts' visual and textual content to approximate the semantic representations of a social account. A modified listwise learning to rank model was learnt to predict ranking scores for the given brands and micro-influencers. Afterward, we utilized the learnt scoring function to recommend micro-influencers for brand promotion. The extensive evaluation demonstrated the advantage of our proposed method compared with the state-of-the-art methods.

In summary, the contributions of our work are as follows:

- We design a novel social account history pooling method, which can leverage social media open data to represent brands and micro-influencers.
- We propose a modified listwise learning to rank model, which successfully predict ranking scores for the given brand and micro-influencers. Extensive experiments validate the effectiveness of our proposed model.
- We collected and organized a brand-micro-influencer dataset[1], which can greatly benefit the future researchers in this area.

2 RELATED WORK

2.1 Influencer Marketing

With the rapid growth of the Internet, millions of users publish more and more personal information on social networks, which makes social media marketing industry flourishing [39]. Moreover, it is constantly soliciting a lot of researches across different disciplines. Especially, works focused on content personalization and deliverance were proposed. Gelli *et al.* [12] proposed a framework that used visual sentiment features together with novel context features to predict a popularity score of social images. Following that, they further proposed a tailored content-based learning-to-rank system to discover content for a target brand in [13]; Mazloom *et al.* [25] utilized engagement parameters for predicting brand-related user post popularity.

[1] https://github.com/gantian/ACMMM-2019-influencer

However, great amount of irrelevant content has destroyed social media users' trust in sponsored search results and online advertisement [15]. To address the distrust issues mentioned above, marketers often seek help from social media influencers, and this marketing method is termed as influencer marketing. Influencer marketing is the "art and science of engaging people who are influential online to share brand messages with their audiences in the form of sponsored content" [14, 18, 34]. In general, influencer marketing derives its value from three sources: *social reaches*, *original content*, and *consumer trust* [29]. Unlike the traditional way of looking only at companies, consumers tend to look at fellow consumers to inform their purchasing decisions. Therefore, influencer marketing is to identify the individuals who have influence over potential customers, and orient marketing activities around these influencers [33].

The top challenge of Influencer marketing is how to find out right influencers for diverse brands. To help marketers find the most right social media micro-influencers at a large scale. Li *et al.* [21] and Segev *et al.* [31] proposed a method for measuring influence for influencers respectively, but both of their work rank influencers without considering brands. Aleksandr *et al.* [8] presented an AI-driven social multimedia influencer discovery marketplace, called SoMin. However, this work only proposed a technical demonstration rather than concrete methods.

2.2 Learning to Rank

Ranking is an essential problem for information retrieval, and has also received extensive attention from the academic research community [20, 22]. Among the common ranking algorithms, learning to rank is a class of techniques that apply supervised machine learning to solve ranking problems.

Learning to rank algorithms can be categorized into three groups by different input spaces, output spaces, hypotheses and loss functions: the pointwise, pairwise, and listwise approach [22]. Taking document retrieval as an example, the ranking task is performed to sort the documents based on its relevance score with respect to a given query. *Pointwise* methods take the feature vector of a document as the input and predict the relevance degree of the document. *Pairwise* methods take a pair of documents as the input and output the relative order between them [22].

But both in pointwise methods and pairwise methods, the group structure of ranking is ignored [20]. *Listwise* methods operate on a group of documents, and predict their relevance degrees or their permutation [36]. For instance, ListNet [2] is one of the first listwise models which uses a loss function defined as cross entropy between parameterized probability distributions of the ground truth and the predicted result. Unfortunately, listwise methods suffer from high computation complexity in model training. To address the problem above, Cao *et al.* [2] used a Top-One approach, which clusters the permutations by top one object. However, ignoring the rank information of partial sequences might lead to ineffective learning. Therefore, Luo *et al.* [23] proposed a stochastic ListNet approach, which samples a small set of object lists, and train the Top-k model based on this small set instead of the full set of permutation classes. But breaking a full object list into partial sequence samples will get a large amount of short object list, which is also a challenge to computation cost.

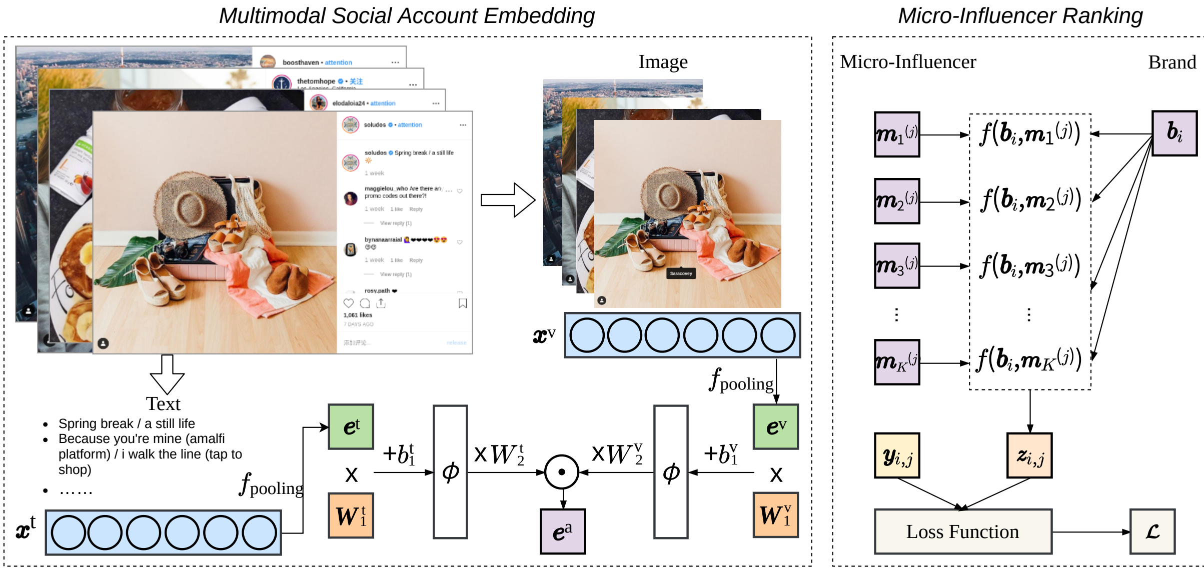

Figure 2: An overview of the proposed framework for micro-influencer ranking.

3 PROPOSED METHOD

Representing brands and micro-influencers is the foundation for recommending micro-influencers to brands. However, there is no standard way to conduct the representation. Therefore, we first proposed a novel multimodal social account embedding method, which exploits social network historical posts to represent brands and micro-influencers. After obtaining the multimodal social account representation, we proposed a modified listwise learning-to-rank model, and then used this model to predict ranking scores for the given brands and micro-influencers. Afterward, we utilized the learnt scoring function to recommend micro-influencers for brand promotion. We illustrate the proposed framework in Figure 2.

3.1 Notations and Problem Formulations

We indicate $\mathbf{B} = \{b_1, b_2, ..., b_{|B|}\}$, and $\mathbf{M} = \{m_1, m_2, ..., m_{|M|}\}$ as the set of brand and micro-influencer list, respectively. Both of them are social media accounts, consisting of posts with visual and textual information. We further define a micro-influencer m as a *positive* example for brand b, if m has posted an advertisement for b. Based on this definition, for every brand b_x, we use $\mathbf{MicroInf}_x^+$, and $\mathbf{MicroInf}_x^-$ to respectively denote its *positive* and *negative* examples, and $\mathbf{MicroInf}_x^+ \bigcup \mathbf{MicroInf}_x^- = \mathbf{M}$.

The goal of our problem is to learn a ranking score function $f(b, m)$ such that for every b_x,

$$f(b_x, m_i) > f(b_x, m_j), \tag{1}$$

where $m_i \in \mathbf{MicroInf}_x^+$ and $m_j \in \mathbf{MicroInf}_x^-$.

3.2 Multi-modal Social Account Representation

3.2.1 Social Account History Pooling. For each social account (brand and micro-influencer), we analyze its recent N_{hist} posts (consisting of both images and texts). We utilized pretrained CNN [32] to extract visual features (with the dimension of d_{v}), and Word2Vec [27] to extract textual features (with the dimension of d_{t}). We denoted the extracted visual and textual features of a social account as $\boldsymbol{x}^{\mathrm{v}} \in \mathbb{R}^{N_{\mathrm{hist}} \times d_{\mathrm{v}}}$, and $\boldsymbol{x}^{\mathrm{t}} \in \mathbb{R}^{N_{\mathrm{hist}} \times d_{\mathrm{t}}}$, respectively.

We applied pooling on extracted deep features for a compact representation. Average pooling is a commonly used sub-sampling method. However, it degrades the performance by losing crucial information in strong activation values [37]. Therefore, we proposed a weighted history pooling method, which subsamples the social account's historical information based on its statistics. Specifically, a weighted pooling f_{weighted} is defined by multiplying a weight based on the standard deviation and the mean of the feature value on an average pooling function f_{ave}:

$$f_{\mathrm{ave}}(\boldsymbol{x}) = \frac{1}{N} \sum_{i=1}^{N_{\mathrm{hist}}} \boldsymbol{x}_{i,:} \tag{2}$$

$$f_{\mathrm{weighted}}(\boldsymbol{x}) = \boldsymbol{w} \odot f_{\mathrm{ave}}(\boldsymbol{x}), \tag{3}$$

where $\odot$ is the element-wise multiplication, and each weight w_j in $\boldsymbol{w}$ is defined as

$$w_j = \exp(-\gamma \frac{s(\boldsymbol{x}_{:,j})}{\mu(\boldsymbol{x}_{:,j})}), \tag{4}$$

where γ is a scaling factor, $s(\cdot)$ is the standard deviation, and $\mu(\cdot)$ is the mean.

We further regarded image and text as two different views to characterize an account. Therefore, instead of conducting post-level analysis, we separated visual and textual features and create their corresponding representations $\boldsymbol{e}^{\mathrm{v}}$ and $\boldsymbol{e}^{\mathrm{t}}$:

$$\boldsymbol{e}^{\mathrm{v}} = f_{\mathrm{pooling}}(\boldsymbol{x}^{\mathrm{v}}), \tag{5}$$

$$\boldsymbol{e}^{\mathrm{t}} = f_{\mathrm{pooling}}(\boldsymbol{x}^{\mathrm{t}}), \tag{6}$$

where $\boldsymbol{e}^v \in \mathbb{R}^{d_{\mathrm{v}}}$, $\boldsymbol{e}^t \in \mathbb{R}^{d_{\mathrm{t}}}$, and pooling $\in$ {ave, weighted}.

Figure 3 illustrates the social account history pooling process.

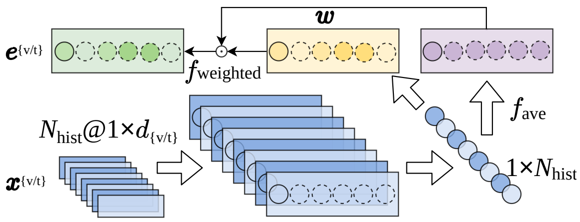

Figure 3: Illustration of the social account history pooling.

3.2.2 Multimodal Social Account Embedding. Though we compress multiple historical data into one representation (i.e., e^v and e^t), a simple concatenation of these two representations tends to be high-dimensional and with imbalanced dimension ($d_v \gg d_t$). To tackle this issue, we adopted a low-rank bilinear pooling method [19] to fuse visual and textual information. It has been successfully applied in various visual tasks like object recognition and segmentation. Specifically, we applied a linear transformation followed by a non-linear activation on each feature to reduce the difference between the size of two feature dimensions. We further projected the joint representations into a given-size output vector. At last, an inner product was applied to capture the multi-modal interaction information.

Formally, the final social account representation e^a is defined as:

$$e^a = \langle \phi(e^t W_1^t + b_1^t) W_2^t, \phi(e^v W_1^v + b_1^v) W_2^v \rangle, \quad (7)$$

where $\langle \cdot, \cdot \rangle$ is the inner product, ϕ denotes a non-linear activation function, $W_1^t \in \mathbb{R}^{d_t \times d_{h1}}$, $b_1^t \in \mathbb{R}^{d_{h1}}$, $W_2^t \in \mathbb{R}^{d_{h1} \times d_a}$, $W_1^v \in \mathbb{R}^{d_v \times d_{h2}}$, $b_1^v \in \mathbb{R}^{d_{h2}}$, $W_2^v \in \mathbb{R}^{d_{h2} \times d_a}$, $e^a \in \mathbb{R}^{d_a}$, d_{h1} and d_{h2} are the length of hidden state vectors, and d_a is the length of the final social account representation.

3.3 Micro-influencer Ranking

3.3.1 Micro-Influncer Competence Score. Unlike celebrities with ready-made audiences who are difficult to engage, micro-influencers are more approachable and affordable. They also enjoy the advantage of better engagement with their audiences than their high-profile counterparts do. However, finding out suitable micro-influencers always involves tedious work by searching millions of posts for any keyword, hashtag or mention, with dozens of filters. Therefore, instead of using only yes or no to represent micro-influencers' relatedness to a brand, we designed a competence score for micro-influencers with respect to every given brand. Specifically, we integrated *engagement* and *relatedness* into a **competence score** of each micro-influencer m_i with respect to brand b_j as:

$$cs(m_i, b_j) = \alpha \, \text{Engagement}(m_i) + (1-\alpha) \, \text{SIM}(m_i, b_j). \quad (8)$$

Engagement was defined as the average number of likes and comments for the posts that m_i used to advertise b_j:

$$\text{Engagement}(m_i) = \frac{\text{AVE}(\#(\text{likes}) + \#(\text{comments}))}{\#(\text{followers})}, \quad (9)$$

where AVE$(\cdot)$ is the average operation, and $\#(\cdot)$ counts the number of items. Meanwhile, the similarity function between two account representation was defined as:

$$\text{SIM}(m_i, b_j) = \frac{e^a(m_i) \cdot e^a(b_j)}{||e^a(m_i)|| \cdot ||e^a(b_j)||} \quad (10)$$

3.3.2 Learning to Rank. Learning to rank refers to machine learning techniques for training the model in a ranking task. In learning to rank algorithms, a ranking function is learnt to assign score values to a collection of documents with respect to given queries. As reviewed previously, listwise approaches directly look at the entire list of documents and try to come up with the optimal ordering for it. One of the most well-known listwise learning methods is ListNet [2]. With the formulation described in the previous sections, we can formulate our micro-influencer recommendation problem with the ListNet framework.

Formally, given the brand list B and micro-influencer list M defined in Sec. 3.1, each brand b_i can be associated with a list of micro-influencer ranking scores y_i as the label:

$$y_i = \big(cs(b_i, m_1), cs(b_i, m_2), \ldots, cs(b_i, m_{|M|})\big). \quad (11)$$

The score y_i represents the degree of competence of posting advertisement with each micro-influencer for brand b_i. The assumption is that the higher the score is observed for b_i and m_j, the stronger relevance exists between them. We then created a ranking score function $f(b, m)$ such that for every b_i,

$$f(b_i, m_x) > f(b_j, m_y), \quad (12)$$

where $m_x \in \mathbf{MicroInf}_i^+$ and $m_y \in \mathbf{MicroInf}_i^-$. Therefore, for each brand b_i, we can obtain a list of scores as,

$$z_i = \big(f(b_i, m_1), f(b_i, m_2), \ldots, f(b_i, m_{|M|})\big). \quad (13)$$

At last, the goal of our problem was to learn the ranking score function $f(b, m)$ by minimizing of the total losses:

$$\sum_{i=1}^{|B|} \mathcal{L}(y_i, z_i), \quad (14)$$

where $\mathcal{L}$ is a listwise loss function.

In ListNet, the loss function is defined as the cross entropy between two distributions: the distribution of human-labeled scores (the competence score in our case), and the probability of an object of being ranked on the top of all objects. One shortcoming of this approach is that it learns the rank information of the full list, but ignores the rank information of partial sequences, which may lead to ineffective learning [23]. For example, if the list contains an object with a much higher score than others, the learning would be largely dominated by the highest score while neglecting the ranking information conveyed by other objects.

3.3.3 $\mathcal{K}$-Buckets Sampling Strategy. In order to address the issues raised above, we broke the full ranking list into multiple length-K partial sequence samples. However, the number of length-K permutation sequences are too large to handle. Therefore, we proposed a novel $\mathcal{K}$-Buckets sampling method to select the samples. The bucket defines a certain pattern of positive-negative examples ratio in a sample, and the micro-influencer examples were then filled into buckets to create the length-K samples. Fig. 4 illustrates the steps of the sampling strategy.

Bucket Creation: we created K buckets such that each bucket $buck_i^K$ will be filled with length-K samples consisting of i positive examples and $K - i$ negative examples. Suppose the average number of brands' positive micro-influencer examples is A, with bucket's definition, the upper bound of the total number of samples in all-positive-example bucket, i.e., $buck_K^K$, is C_A^K. The total number

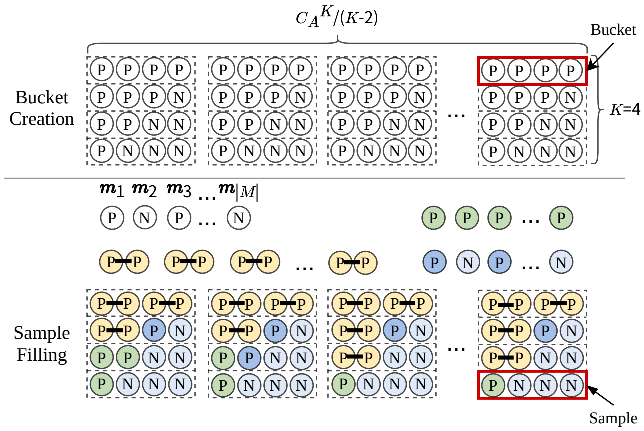

Figure 4: Illustration of the $\mathcal{K}$-Buskets sampling strategy.

of brand-micro-influencer samples is $C_A^1 \cdot \frac{(|M|-A)!}{(|M|-A-K+1)!} + C_A^2 \cdot \frac{(|M|-A)!}{(|M|-A-K+2)!} + ... + C_A^K) \cdot |B|$, which is computational prohibitive when $K > 2$. For example, using the setting in our experiment with $B = 286$, $M = 3146$, and $A = 11$, the number of samples is around 9 million for $K = 2$, and 30 billion for $K = 3$. Therefore, we constrain the capacity of the buckets with the following rules: 1) the number of samples in each bucket is the same; and 2) the number of samples in bucket $buck_K^K$ is set as $C_A^K/(K-2)$. Consequently, the capacity of K buckets is $N_{\text{sampling}} = (C_A^K * K)/(K-2)$.
Sample Filling: Since the goal of our learning is to separate positive and negative examples, it would be better to ensure the occurrance of every positive examples. We first picked every positive examples into the buckets starting from $buck_1$. After that, we enumerated the combination of any two positive examples and filled them into the free slots in $buck_2$ to $buck_K$. At last, random sampling was applied to the rest free slots with the restriction that no identical examples in one bucket and $buck_i$ should contain i positive examples.

3.3.4 Top-1-over-K Probability. With the notation defined above, suppose $(m_1^{(j)}, m_2^{(j)}, \ldots, m_K^{(j)})$ is the j-th sequence after our $\mathcal{K}$-Buckets sampling, the ground truth for each sequence is modified from Equation (11) into:

$$\hat{y}_{i,j} = \sigma\big(cs(b_i, m_1^{(j)}), cs(b_i, m_2^{(j)}), \ldots, cs(b_i, m_K^{(j)})\big). \quad (15)$$

Similarity, the list of predicted score is modified into:

$$\hat{z}_{i,j} = \sigma\big(f(b_i, m_1^{(j)}), f(b_i, m_2^{(j)}), \ldots, f(b_i, m_K^{(j)})\big), \quad (16)$$

where σ is the softmax function.

3.3.5 Loss Functions. The final loss function becomes:

$$\mathcal{L} = -\frac{1}{|B|}\sum_{i=1}^{|B|}\sum_{j=1}^{N_{\text{sampling}}} \hat{y}_{i,j}\log(\hat{z}_{i,j}) + \lambda||\theta||_1, \quad (17)$$

where θ is the set of all the weights of the model, and λ controls the importance of the regularization terms.

4 DATASET CONSTRUCTION

There is no open dataset available for micro-influencer recommendation. As one of the best social media platforms for engagement, Instagram, specializing photos and micro-videos, is selected as our data source. With over 1 billion monthly active users (announced by their company in June 2018), it is currently one of the most popular social media platforms for influencer marketing. Therefore, we crawled publicly available data on the Instagram social platform through the site's official API, and built a Brand-Micro Influencer dataset. It consists of 360 brands and 3,748 micro-influencers, where each brand and each micro-influencer is an Instagram account.

We started with the dataset from the work [13], which consists of 900 brands from 14 categories[2]. We removed three categories (*Furnishing*, *Finance*, and *Energy*) which contain less than 50 brands, and replaced *Fashion* with *clothing* and *shoes*, to finally obtain 12 categories. Then we collected trending Instagram hashtags from social media hashtag analysis websites[3], where each hashtag on these websites is associated with a category label. Therefore, the collected hashtags were used as seeds to further crawl the brand account. At last, we selected 30 brand accounts for each category, resulting in 360 brand accounts in total.

With the brand list, we crawled the latest 1,000 posts of each brand account, and consider the accounts mentioned within these posts as the candidates of micro-influencers. We further crawled these candidates' profile pages to retrieve their biographies and the number of followers. From these candidates, we selected the accounts under the criterion that the number of followers is between 5,000 to 100,000, and removed the non-English accounts. In this way, we paired each brand with around 11 micro-influencers. Note that there are a small portion (around 10%) of micro-influencers belonging to multiple brands.

At last, given the brand and micro-influencer list, we crawled their recent 50 posts (with their visual information, textual information, the number of comments, and the number of likes), and their profile (the number of followers, and the bio description) and constructed our dataset.

5 EXPERIMENTS

5.1 Experimental Setup

We splited our data set into training set and testing set, where the training set contains 286 brands and the testing set contains 74 brands. Each brand has about 11 micro influencers. Both training set and testing set have all kinds of brands belonged to 12 categories.

Throughout the training we initialized all of the neural network parameters with the uniform distribution between -0.1 and 0.1. We used stochastic gradient descent with a learning rate of 0.001 and a decay of 0.9 every epoch. We trained our model for thirty epochs and used a batch size of 64. In our model, we set the scaling factor γ = 1/3, α=0.5 and N_{hist}=50. The length of our text representation $\boldsymbol{e}^{\text{t}}$, and image representation $\boldsymbol{e}^{\text{v}}$ are 300 and 25,088, respectively. The hidden layer length of text (d_{h1}) and image (d_{h2}) are 300 and 4096, respectively. Social account representation length d_{a} is 512. We used leaky ReLU as the activation function ϕ. A dropout layer with rate =0.5, and L1 regularization with a regularization rate λ of 0.001 were used. Besides, we used local response normalization on our features.

[2]Alcohol, Airlines, Auto, Fashion, Food, Furnishing, Electron., Nonprof., Jewellery, Finance, Services, Entertain., Energy, and Beverages.
[3]E.g., https://www.hashtagsforlikes.co, https://hashtagify.me.

5.2 Baselines

In our dataset, only around 10% micro-influencers interact with more than one brand, which will be discarded in collaborative filtering based methods (with less than five interactions). Therefore, collaborative filtering based models like NCF [16], AutoRec [30], and SVDFeature [3] are not designed for our problem since they rely on multiple user-item interactions [13]. In order to demonstrate the effectiveness of our proposed method, we employ the following methods as baselines:

- **RAND**: we generated a random score for each micro-influencer.
- **SBR**: we concatenated text features and image features as account representation, and used cosine similarity scores of brand-micro-influencer pairs to rank, which is a simulation of content-based ranking methods.
- **MIR(k)**: our proposed micro-influencer ranking methods with different bucket size k, and MIR is actually a pairwise method when k=2 and a listwise method when k>2.

The following variants of MIR are designed for comparison:

- **MIR$^{\text{v-only/t-only}}$**: MIR$^{\text{v-only}}$ and MIR$^{\text{t-only}}$ are variants of MIR with visual and textual feature only, respectively. These variants were designed to evaluate the usage of different modalities.
- **MIR-concat/ap/wap**: these three methods differ how we represent multimodal social account information, where *concat* simply concatenate visual and textual features and is followed by a two-layer fully connected neural network; *ap* uses the average pooling to fuse social account's history; and *wap* uses the weighted average pooling.

5.3 Evaluation Metrics

We adopted AUC, cAUC, Rec@k, and MedR to evaluate the performance.

- **AUC** is the probability that a positive example's ranking score is higher than a negative example's ranking score;
- **cAUC** is the AUC where the positive example and the negative example belong to a same category;
- **Rec@k** is the fraction of positive examples that have been recommended in top k over the total amount of positive examples;
- **MedR** is the median position of the first positive example. The lower MedR is, the better the performance.

It is to note that we discarded the metric precision@10 and precision@50 because in our dataset, we have only 11 positive micro-influencers for each brand in average, thus making precision@10 or precision@50 less meaningful.

5.4 Experimental Results and Discussion

Table 1 shows the comparison between our proposed MIR(k) with the baselines. From the table, we can see that, our proposed MIR outperforms all the baselines. Also, the AUCs are consistently lower than the cAUCs. This is mainly because that the metric cAUC focuses on evaluating the ability of differentiating brands within the same category, which is a tougher task than that in all categories. For the comparison between our proposed MIR with different k values, we can see that the AUC and cAUC do not change much, however, a notable improvement for Rec@10, Rec@50, and MedR was observed. Specifically, compared to MIR(2), MIR(3) increased 26.2% in Rec@10, 8.5% in Rec@50, and 3 positions in MedR. Rec@10, Rec@50, and MedR reflect how the methods perform for the top ranked micro-influencers. This confirms our argument that listwise learning method (k>2) can lead to better performance than the pairwise learning (k=2).

Table 1: Comparison of MIR(k) with the baselines. The MIR(k)s use both visual and textual information, and use the weighted average pooling for social account history pooling.

Method	AUC	cAUC	Rec@10	Rec@50	MedR
RAND	0.494	0.489	0.005	0.054	54
SBR	0.681	0.593	0.091	0.251	15
MIR(k=2)	0.841	0.672	0.103	0.390	9
MIR(k=3)	0.845	0.673	0.130	0.423	**6**
MIR(k=4)	**0.849**	**0.675**	**0.135**	**0.428**	**6**

Table 2: Comparison of MIR variants (visual only, textual only, feature concatenation, average pooling, and weighted average pooling).

Method	AUC	cAUC	Rec@10	Rec@50	MedR
MIR$^{\text{v-only}}$	0.829	0.659	0.095	0.382	8
MIR$^{\text{t-only}}$	0.531	0.517	0.014	0.076	57
MIR-concat	0.691	0.592	0.034	0.153	26
MIR-ap	**0.845**	**0.674**	0.096	0.415	8
MIR-wap	**0.845**	0.673	**0.130**	**0.423**	**6**

Table 2 lists the comparison between different MIR variants. All these variants are with k = 3 for a fair comparison. For the comparison between using MIR with visual-only and textual-only data, the MIR$^{\text{v-only}}$ is significantly better than MIR$^{\text{t-only}}$. This is not supersing because different from Facebook and Twitter, Instagram is a visual-based social media platform with very little focus on text. Nevertheless, textual information is still useful since the full usage of both visual and textual information consistently outperforms that of visual-only (see the comparison between MIR-wap and MIR$^{\text{v-only}}$). We also conducted a comparison between different account representation methods (i.e., *concat*, *ap*, and *wap*). We can see that simple concatenation of visual and textual features does not help (with lower performance than MIR$^{\text{v-only}}$). Though AUC and cAUC have the similar performance between average pooling and weighted average pooling methods, yet, MIR-wap is better in all three Rec@10, Rec@50, and MedR. This indicates that the weighted average pooling helps the model to recommend better top ranked micro-influencers.

In order to deep dive into how the proposed method performs on the data in different categories, we report the details of the results on our 12 categories in Table 3. We can see that *makeup* has the best performance in AUC and *entertainment* performs best in cAUC. One interesting thing to note is that the results on category *auto* and *food*, have a larger gap between their cAUC and AUCs compared

Table 3: Performance for different brand categories.

Category	AUC	cAUC	Rec@10	Rec@50	MedR
airline	0.816	0.657	0.042	0.345	21
auto	0.862	0.541	0.133	0.508	6
clothing	0.885	0.717	0.082	0.426	5
drink	0.794	0.694	0.073	0.306	8
electronics	0.747	0.658	0.054	0.219	9
entertainment	0.873	**0.738**	0.105	0.432	6
food	0.812	0.554	0.082	0.424	9
jewelry	0.758	0.692	0.070	0.276	4
makeup	**0.928**	0.675	0.159	**0.723**	3
nonprofit	0.863	0.697	**0.184**	0.450	**1**
shoes	0.920	0.702	0.117	0.658	4
services	0.851	0.726	0.069	0.254	10

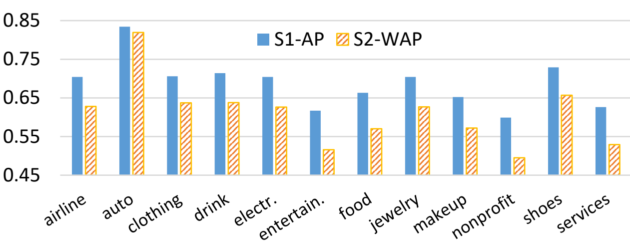

Figure 5: Brand categorical self-similarity. S1-AP denotes the self-similarity with average history pooling, and S2-WAP denotes the self-similarity with weighted history pooling.

to other categories. To gain an insight into this issue, we further calculated a brand categorical self-similarity between brands within the same category and plotted it in Figure 5. The similarity is defined as the cosine similarity between the concatenation of visual and textual features. As we can see from Figure 5, *auto* category has a much higher categorical self-similarity than other categories, which is the main reason for its poor performance. The *food* category has much lower self-similarity. However, after we analyzed the predicted ranking results, we found out that our model predicts higher score for a few food bloggers who share all kinds of food in Instagram. Though the incorrect prediction downgrades its cAUC performance, we believe that these food bloggers are still potential micro-influencers for food brands.

We have also reported a comparison of categorical self-similarity between the average pooling method ('S1-AP' in Figure 5) and the weighted average pooling method ('S2-WAP' in Figure 5). We can see that the weighted average pooling method helps to reduce the brand feature categorical self-similarity.

5.5 Case Study

To have an intuitive understanding of our proposed method, we present examples of micro-influencer recommendation using the method described in Section 3. As shown in Figure 6, each row represents a brand, the top one positive example, and the top two negative examples recommended by our method. Brands (a) to (d) come from category *makeup*, *clothing*, *auto*, and *drink*, respectively. We selected four representative images from each brand/micro-influencer social account, followed by its category, biography description and the images' corresponding textual descriptions.

From the correctly predicted positive examples we can see that, generally, brands and their corresponding micro-influencers are closely related to each other in their own field. To be specific, the *makeup* brand has a positive example of "Makeup and eyebrow specialist" micro-influencer, and the positive example of *auto* is a "Bentley Factory Driver". Moreover, these micro-influencers tend to post images/texts closely related to the posts of the brand accounts.

For the negative examples predicted by our model, we have the following observations: 1) the post information in *makeup* is quite unique (with face, eyes, or lips occupying the whole image), however, it is rather difficult to differentiate micro-influencers within this category. This may explain why *makeup* has 0.928 for its AUC while only 0.675 in its cAUC (cf. Table 3); 2) the model will predict micro-influencers out of their actual category, however, these negative examples belong to category closely related to the brand, such as the *makeup* in (b). And 3) some of the micro-influencers are labelled with an unexpected category, however, it is still explainable. For example the Negative Example 1 in (d) shows photos of all kinds of food, the account is actually promoting a camera by showing high-quality photos shooted by that camera, thus it falls into *electronics* category. Similarly, the Negative Example 2 in (d) always displays dedicated food images when s/he travels outside, however, s/he has posted an advertisement for a car with an image showing s/he drives the car for travelling. These information are hardly to infer, thus the textual description may be helpful in such cases.

6 CONCLUSIONS AND FUTURE WORKS

In this work, we proposed a data-driven micro-influencer ranking scheme to solve the essential question of finding out the right micro-influencer. We have also developed a new Instagram brand micro-influencer dataset, which can benefit future researchers in this area. The extensive evaluation demonstrates the advantage of our proposed method compared with the state-of-the-art methods.

Several research topics are open for future investigation. During result analysis, we found out certain micro-influencers can give advertisement across different categories. We would like to consider whether cross-category information can help improve the performance. In addition, though textual information did not perform well in our experiment, we still believe that textual information is useful when visual information did not discriminate different micro-influencers. Also, we want to investigate micro-influencer marketing in other social network platforms like Twitter and Facebook.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China, No.: 61772310, No.: 61702300, No.: 61702302, No.: 61802231, No.:U1836216; the Project of Thousand Youth Talents 2016; the Shandong Provincial Natural Science and Foundation, No.: ZR2019JQ23, No.: ZR2019QF001; the Future Talents Research Funds of Shandong University, No.: 2018WLJH63; the Fundamental Research Funds of Shandong University (No. 2017HW001).

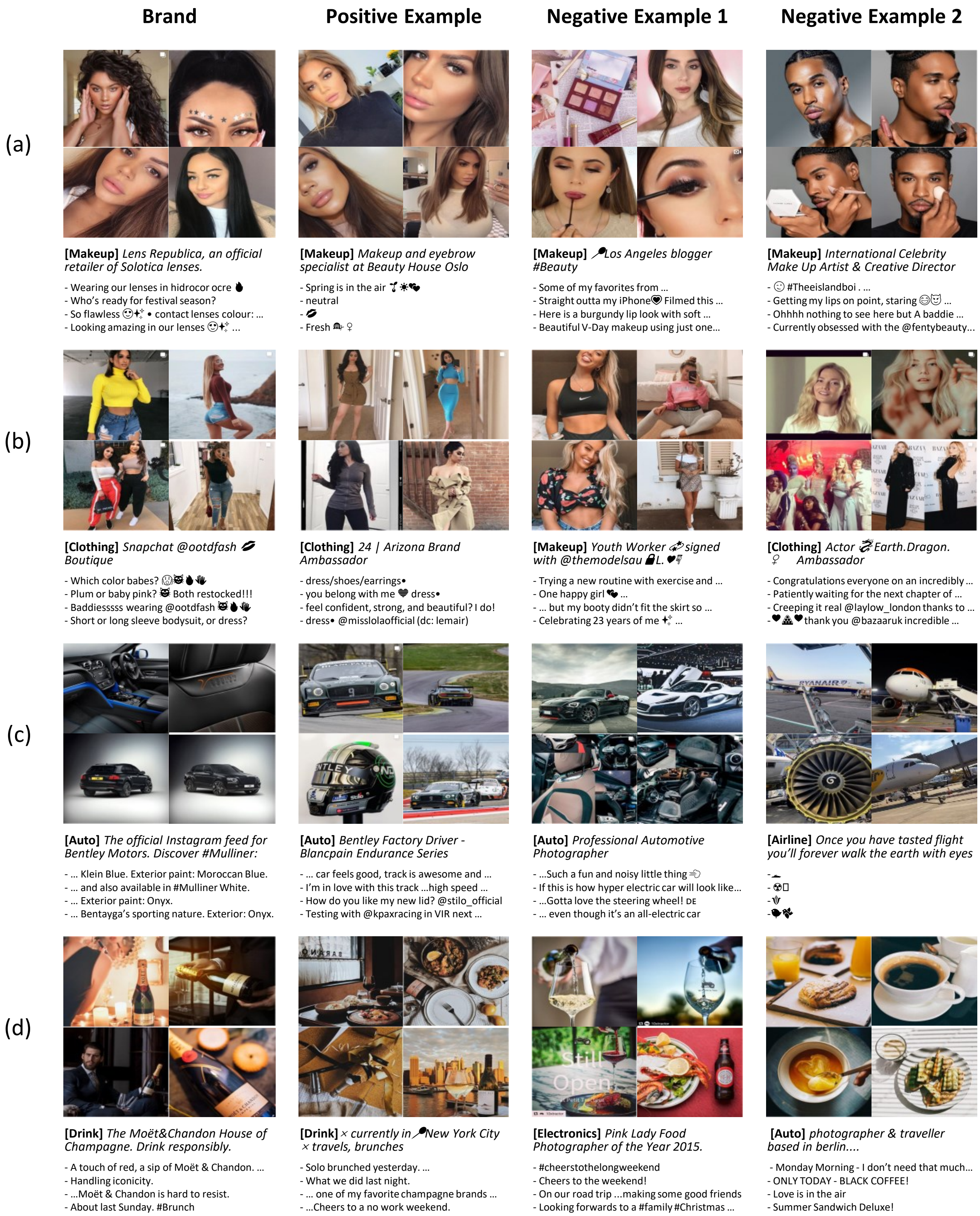

Figure 6: Case study for brands with their positive/negative examples.

REFERENCES

[1] Christy Ashley and Tracy Tuten. 2015. Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement. *Psychology & Marketing* 32, 1 (2015), 15–27.

[2] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to rank: from pairwise approach to listwise approach. In *Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML), Corvallis, Oregon, USA*. 129–136.

[3] Tianqi Chen, Weinan Zhang, Qiuxia Lu, Kailong Chen, Zhao Zheng, and Yong Yu. 2012. SVDFeature: a toolkit for feature-based collaborative filtering. *Journal of Machine Learning Research* 13, Dec (2012), 3619–3622.

[4] Eun-Kyong Choi, Deborah Fowler, Ben Goh, and Jingxue Yuan. 2016. Social media marketing: applying the uses and gratifications theory in the hotel industry. *Journal of Hospitality Marketing & Management* 25, 7 (2016), 771–796.

[5] Marijke De Veirman, Veroline Cauberghe, and Liselot Hudders. 2017. Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *International Journal of Advertising* 36, 5 (2017), 798–828.

[6] Laurence Dessart, Cleopatra Veloutsou, and Anna Morgan-Thomas. 2015. Consumer engagement in online brand communities: a social media perspective. *Journal of Product & Brand Management* 24, 1 (2015), 28–42.

[7] Nathaniel J Evans, Joe Phua, Jay Lim, and Hyoyeun Jun. 2017. Disclosing Instagram influencer advertising: The effects of disclosure language on advertising recognition, attitudes, and behavioral intent. *Journal of Interactive Advertising* 17, 2 (2017), 138–149.

[8] Aleksandr Farseev, Kirill Lepikhin, Kenny Powar, Eu Khoon Ang, and Hendrik Schwartz. 2018. SoMin. ai: Social Multimedia Influencer Discovery Marketplace.. In *ACM Multimedia Conference on Multimedia Conference, Demo*.

[9] Reto Felix, Philipp A Rauschnabel, and Chris Hinsch. 2017. Elements of strategic social media marketing: A holistic framework. *Journal of Business Research* 70 (2017), 118–126.

[10] Tian Gan, Yongkang Wong, Bappaditya Mandal, Vijay Chandrasekhar, and Mohan S. Kankanhalli. 2015. Multi-sensor Self-Quantification of Presentations. In *ACMMM*. 601–610.

[11] Tian Gan, Yongkang Wong, Daqing Zhang, and Mohan S. Kankanhalli. 2013. Temporal encoded F-formation system for social interaction detection. In *Proceedings of ACM International Conference on Multimedia*. 937–946.

[12] Francesco Gelli, Tiberio Uricchio, Marco Bertini, Alberto Del Bimbo, and Shih-Fu Chang. 2015. Image popularity prediction in social media using sentiment and context features. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 907–910.

[13] Francesco Gelli, Tiberio Uricchio, Xiangnan He, Alberto Del Bimbo, and Tat-Seng Chua. 2018. Beyond the Product: Discovering Image Posts for Brands in Social Media. In *ACM Multimedia Conference on Multimedia Conference*. 465–473.

[14] Morgan Glucksman. 2017. The rise of social media influencer marketing on lifestyle branding: A case study of Lucie Fink. *Elon Journal of Undergraduate Research in Communications* 8, 2 (2017), 77–87.

[15] Khim-Yong Goh, Cheng-Suang Heng, and Zhijie Lin. 2013. Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content. *Information Systems Research* 24, 1 (2013), 88–107.

[16] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 173–182.

[17] Linda D Hollebeek, Mark S Glynn, and Roderick J Brodie. 2014. Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of interactive marketing* 28, 2 (2014), 149–165.

[18] Linus Juhlin and Miretta Soini. 2018. How do influencer marketers affect brand associations?: a semiotic Instagram study in the sports fashion industry.

[19] Jin-Hwa Kim, Kyoung Woon On, Woosang Lim, Jeonghee Kim, Jung-Woo Ha, and Byoung-Tak Zhang. 2017. Hadamard Product for Low-rank Bilinear Pooling. In *International Conference on Learning Representations, ICLR, Toulon, France, Conference Track Proceedings*.

[20] Hang Li. 2011. A short introduction to learning to rank. *IEICE TRANSACTIONS on Information and Systems* 94, 10 (2011), 1854–1862.

[21] Yung-Ming Li, Cheng-Yang Lai, and Ching-Wen Chen. 2011. Discovering influencers for marketing in the blogosphere. *Information Sciences* 181, 23 (2011), 5143–5157.

[22] Tie-Yan Liu et al. 2009. Learning to rank for information retrieval. *Foundations and Trends® in Information Retrieval* 3, 3 (2009), 225–331.

[23] Tianyi Luo, Dong Wang, Rong Liu, and Yiqiao Pan. 2015. Stochastic Top-k ListNet. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 676–684.

[24] Rakesh Mallipeddi, Subodha Kumar, Chelliah Sriskandarajah, and Yunxia Zhu. 2018. A Framework for Analyzing Influencer Marketing in Social Networks: Selection and Scheduling of Influencers. *Fox School of Business Research Paper* 18-042 (2018).

[25] Masoud Mazloom, Robert Rietveld, Stevan Rudinac, Marcel Worring, and Willemijn van Dolen. 2016. Multimodal Popularity Prediction of Brand-related Social Media Posts. In *Proceedings of the ACM Conference on Multimedia Conference*. 197–201.

[26] Nina Michaelidou, Nikoletta Theofania Siamagka, and George Christodoulides. 2011. Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands. *Industrial marketing management* 40, 7 (2011), 1153–1159.

[27] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *International Conference on Learning Representations, ICLR, Scottsdale, Arizona, USA, Workshop Track Proceedings*. http://arxiv.org/abs/1301.3781

[28] Stefan Olof Lagrosen and Kerstin Grundén. 2014. Social media marketing in the wellness industry. *The TQM Journal* 26, 3 (2014), 253–260.

[29] Kristy Sammis, Cat Lincoln, and Stefania Pomponi. 2015. *Influencer marketing for dummies*. John Wiley & Sons.

[30] Suvash Sedhain, Aditya Krishna Menon, Scott Sanner, and Lexing Xie. 2015. Autorec: Autoencoders meet collaborative filtering. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 111–112.

[31] Noam Segev, Noam Avigdor, and Eytan Avigdor. 2018. Measuring Influence on Instagram: A Network-Oblivious Approach. In *The International ACM SIGIR Conference on Research & Development in Information Retrieval*. 1009–1012.

[32] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations, ICLR, San Diego, CA, USA, Conference Track Proceedings*. http://arxiv.org/abs/1409.1556

[33] Tracy L Tuten and Michael R Solomon. 2017. *Social media marketing*. Sage.

[34] Chad Witkemper, Choong Hoon Lim, and Adia Waldburger. 2012. Social media and sports marketing: Examining the motivations and constraints of Twitter users. *Sport Marketing Quarterly* 21, 3 (2012).

[35] Steven Woods. 2016 (accessed April 1, 2019). *#Sponsored: The Emergence of Influencer Marketing*. https://trace.tennessee.edu/utk_chanhonoproj/1976/

[36] Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, and Hang Li. 2008. Listwise approach to learning to rank: theory and algorithm. In *Machine Learning, Proceedings of the International Conference (ICML)*. 1192–1199.

[37] Matthew D. Zeiler and Rob Fergus. 2013. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks. In *International Conference on Learning Representations, ICLR, Scottsdale, Arizona, USA, Conference Track Proceedings*.

[38] Xiaoyu Zhu, Tian Gan, Xuemeng Song, and Zhumin Chen. 2017. Sentiment Analysis for Social Sensor. In *Advances in Multimedia Information Processing, Pacific-Rim Conference on Multimedia*. 893–902.

[39] Nathalie Zietek. 2016. Influencer Marketing: the characteristics and components of fashion influencer marketing.