Multimodal Popularity Prediction of Brand-related Social Media Posts

Mazloom, M.; Rietveld, R.; Rudinac, S.; Worring, M.; van Dolen, W.

DOI
10.1145/2964284.2967210

Publication date
2016

Document Version
Final published version

Published in
MM'16

License
Article 25fa Dutch Copyright Act

Link to publication

Citation for published version (APA):
Mazloom, M., Rietveld, R., Rudinac, S., Worring, M., & van Dolen, W. (2016). Multimodal Popularity Prediction of Brand-related Social Media Posts. In *MM'16: Proceedings of the 2016 ACM on Multimedia Conference : October 15-19, 2016, Amsterdam, The Netherlands* (pp. 197-201). Association for Computing Machinery. https://doi.org/10.1145/2964284.2967210

Multimodal Popularity Prediction of Brand-related Social Media Posts

Masoud Mazloom[†], Robert Rietveld [⋆], Stevan Rudinac [†], Marcel Worring[†⋆], Willemijn van Dolen[⋆]
[†]Informatics Institute, University of Amsterdam [⋆]Amsterdam Business School, University of Amsterdam

ABSTRACT

Brand-related user posts on social networks are growing at a staggering rate, where users express their opinions about brands by sharing multimodal posts. However, while some posts become popular, others are ignored. In this paper, we present an approach for identifying what aspects of posts determine their popularity. We hypothesize that brand-related posts may be popular due to several cues related to factual information, sentiment, vividness and entertainment parameters about the brand. We call the ensemble of cues *engagement parameters*. In our approach, we propose to use these parameters for predicting brand-related user post popularity. Experiments on a collection of fast food brand-related user posts crawled from Instagram show that: visual and textual features are complementary in predicting the popularity of a post; predicting popularity using our proposed *engagement parameters* is more accurate than predicting popularity directly from visual and textual features; and our proposed approach makes it possible to understand what drives post popularity in general as well as isolate the brand specific drivers.

1. INTRODUCTION

More than ever social media, where people are expressing their opinions through posts, is reshaping the business and marketing landscape. It has become an invaluable source for companies to understand the engagement of people with their brand or product. Instagram in particular, with its focus on self-expression and engagement, might provide an answer to the question of which aspects make some posts about brands more popular than other. Some of the information to answer this question is explicitly captured in parameters such as the number of likes. The bulk of information, however, is hidden in the text and visual content of the post. To benefit from the abundance of data in posts for understanding brand-related user post popularity, we need to be able to leverage both the explicit information in a post as well as the visual and textual content.

MM '16, October 15-19, 2016, Amsterdam, Netherlands
© 2016 ACM. ISBN 978-1-4503-3603-1/16/10...$15.00
DOI: http://dx.doi.org/10.1145/2964284.2967210

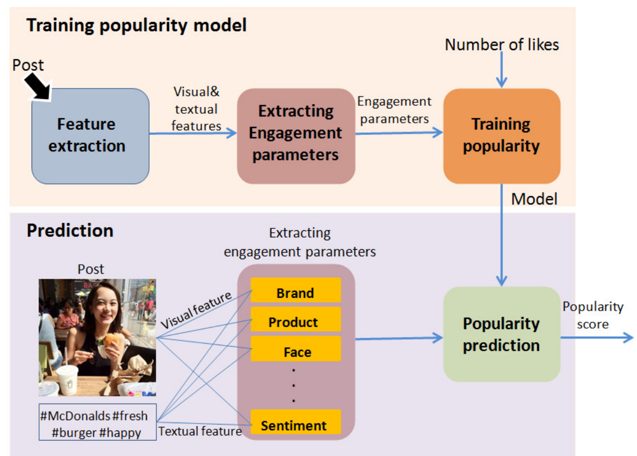

Figure 1: Our proposal for computing popularity of brand-related posts by considering a new layer consisting of *engagement parameters*.

Explicit parameters and textual content have been studied extensively to understand popularity of tweets on Twitter [1, 8] and videos on YouTube [15]. In [8], Hong *et al.* show the importance of retweets for the prediction of message popularity on Twitter. The authors in [1] found a relation between the sentiment and popularity of tweets. Szabo et al. in [15] study the popularity of YouTube videos by analyzing the social cues, comments, and associated tags. All of these works successfully predict popularity using textual data. Visual content, which also holds a lot of information, is not addressed in these methods.

In recent years, various aspects of predicting the popularity of images have been studied by analyzing data from the photo-sharing site Flickr [3, 6, 9, 12]. In [3], Cappallo *et al.* approach popularity prediction as a retrieval problem and learn a ranker that takes into account the distinctive visual cues in popular and unpopular images. Kholsa *et al.* [9] and McParalane *et al.* [12], model image popularity by learning regressors on both visual and textual context. Gelli *et al.* in [6] study the effect of visual sentiment features as well as contextual features. These methods demonstrate the potential of predicting the popularity for image content. Yet, in case of images from a photo-sharing site, the objective information as well as aesthetic quality are likely the determining factors for popularity. For Instagram, with its high portion of selfies and other forms of self-expression, different factors may determine the popularity and hence other or additional algorithms may be more appropriate.

We hypothesize that brand-related user posts may be popular due to several visual and textual cues such as the presence of the logo, products, faces, specific objective content descriptors, and sentiment evoked in the user posts. Therefore, we propose a new brand post popularity detection method by estimating the factual information, sentiment, vividness and entertainment parameters of the brand-related post. These aspects are based on frameworks from the context specific field, i.e. Marketing. The correlation between each of these parameters and the number of likes of each post (see Figure 1) provides a measure of predictive quality. By making the parameters contributing to brand post popularity explicit contributes to understanding the drivers of post popularity as well as increasing the prediction accuracy.

We make the following contributions in this paper:

- We are the first to study automatic multimodal brand popularity prediction on Instagram.
- We introduce a new layer of *engagement parameters* inspired by marketing research.
- We reveal which *engagement parameters* are most influential in making a brand-related user posts popular and how this varies between brands.

2. THE PROPOSED METHOD

In this section we describe our approach to popularity prediction by *engagement parameters*. We first define the parameters which play an important role in marketing. Then, we show how to extract and use them for predicting the popularity of brand-related posts.

2.1 Defining Engagement Parameters

We aim at developing an effective predictive model of post popularity in social media. We hypothesize that brand-related posts may be popular due to several reasons. For example, a marketeer analyzing the popularity of *McDonald's* posts in social media may wonder whether a post is popular because it comes from a famous person, the logo is visible, some brand specific products are present, or that it evokes a certain sentiment. We conjecture that the information needed to answer such questions may be implicitly or explicitly captured in the visual and textual channel of user posts. To study popularity, we first need to identify parameters that are context-specific to Marketing and are known from that field to influence popularity.

We are inspired by recent research in the marketing community for predicting the popularity of a post [4, 14]. In [4], De Vries *et al.* consider information, entertainment and vividness as important sources for predicting brand post popularity. Smith *et al.* in [14] investigate the popularity by identifying four parameters in each post: promotional self-presentation, brand centrality, factually informative of the brand, and brand sentiment. We adapt and combine these parameters to define nine parameters, *engagement parameters*, which potentially impact the popularity of user generated brand related posts. Our *engagement parameters* are the following: *Brand* (brand centrality), *Product* (factual information about the brand), *Sentiment* (brand sentiment), *Image aesthetics* (vividness), *Concepts* (information and entertainment), *Face* (promotional self-presentation). We also combine *Product* and *Face* to produce *Person-Product* and *People-Product* which might reveal whether a brand may be perceived as more individual or social. At the end we consider number of *Followers* as a parameter which has direct effect on popularity. Different from [4, 14], which rely on manual classification for extracting drivers of image popularity, we propose to automatically detect the parameters that drive post popularity and effectively incorporating them into a model for popularity prediction of brand-related posts.

2.2 Extracting Engagement Parameters

To detect the existence of *engagement parameters* in user posts, we analyse their visual and text content. For detecting the presence of a *Brand, Face(s)*, and *Product(s)* in the image of each post, we use the Google Vision API.

Brand: Binary indication of the *Brand* logo presence.

Face(s): The number of detected faces.

Product: In order to assess whether a *Product* is present in an image, we use the following procedure. We select the top 10 labels extracted by *Google Vision API* based on their confidence score. Each of the labels corresponds to an entry in Freebase. A label is considered to be a *Product* of the fast food restaurant if an entry was found in the food or dish category of Freebase. We consider a *Product* to be present in the image if at least one match was found.

Person-Product: Binary indication of the coincidence of exactly one *Face* and *Product*.

People-Product: The value of this parameter is 1 if there are 2 or more *Faces* and *Product* in image.

Sentiment: To detect the sentiment in user posts, we perform late fusion of *Visual-Sentiment* and *Textual-Sentiment*. We determine the *Visual-Sentiment*, similar to [6], using Sentibank detectors [2]. The ontology consists of a collection of 1,200 Adjective-Noun-Pairs (ANP). So we represent each image with a 1,200-dimensional vector in which each dimension shows the probability of the ANP being present in the image. *Textual-Sentiment* is determined by analyzing the text from hashtags, captions, and the comments associated with the images. Due to its effectiveness in extracting sentiment from short informal texts and for consistency with [2], we then deploy SentiStrength [17] to generate the positive (ranging from 1 to 5) and negative (ranging from -1 to -5) sentiment scores.

Image aesthetics: There are in total 42 filters in Instagram which the users can apply before sharing their images. We extract the information about the filters applied to the image from the contextual metadata of the post in to a 42-dimensional binary vector.

Concepts: we extract the convolutional neural network features, initially proposed in [16] and trained to identify those 15,293 ImageNet [5] concept categories, for which at least 200 positive examples are available. Similar to [10, 11], we represent the image of each post by the 15,293-dimensional output of the softmax layer of the network.

Followers: The number of contacts of the given user.

2.3 Post Popularity Measurement

Similar to [6, 9], we treat the post popularity prediction as a ranking problem. As the measure of post popularity we use the number of likes it received. We find that the number of likes follows a power law distribution, where the majority of posts receive little or no likes and the minority of them receive a high number of likes. To deal with the large variation in the number of likes, we apply the log function to make it resemble a Gaussian distribution.

Table 1: Experiment 1: Analysis of different visual and textual features on brand-related post popularity prediction.

	Visual features			Feature fusion	Multimodal fusion
Datasets	***CNN-Pool5***	***ConceptVec-15k***	***Visual-Sentiment***	***Visual-Combi***	***Visual-Textual-Combi***
Category-mix	0.201	0.187	0.224	0.233	*See*
Category-specific	0.220	0.196	0.264	0.270	*below*
	Textual features			**Feature fusion**	**Multimodal fusion**
Datasets	***W2V***	***Textual-Sentiment***	***Term-based***	***Textual-Combi***	***Visual-Textual-Combi***
Category-mix	0.184	0.206	0.191	0.231	0.252
Category-specific	0.217	0.241	0.220	0.250	0,290

In order to predict post popularity, we use L2 regularized L2 loss Support Vector Regression (SVR) from the LIBLINEAR package. We train a popularity model on the training set and use it for predicting the popularity of a post at test time (see Figure 1). To find the optimal value of the regularization parameter C of SVM, we perform 5-fold cross-validation for $C \in \{0.001, 0.1, 1, 10, 100\}$.

3. EXPERIMENTAL SETUP

Dataset: We investigate the effectiveness of our proposed approach for post popularity prediction on an Instagram dataset. As mentioned in [7] the fast food category is one of the most popular categories on Instagram. As we are interested in analyzing brands, we select the six most popular fast food brands based on the annual revenue from 2014, namely: *McDonald's, Burger King, Culver's, Wendy's, Sonic Drive In*, and *Jack In The Box*. We created a dataset by crawling 75,000 user-posts related to these six categories.

In order to explore the effect of the *engagement parameters* on all user posts or in a specific brand category, we define two different settings:

Category-mix: In this setting the whole dataset is used to build a general model for popularity prediction.

Category-specific: For this setting, we perform the training and evaluation independently for each brand category.

In both settings we split the data randomly into two halves, one for training and the other for testing. We average the performance over 10 random splits to ensure consistency of the results.

Baseline features: *Visual features*, As low-level visual features we used the *CNN-Pool5*, the 1024-dimensional features from pooling the last fully connected layer of the Deep Net in [16]. We use also *Concepts*, and *Visual-Sentiment* explained in section 2.2 as high level features. We use three different representations based on the tags added to the posts as *Textual features*: *Term-based*, a sparse binary representation with a length of 22,110, where the length of the vector equals the number of unique tags in the dataset. Only the values of those elements corresponding to the tags of the post are set to one. The *W2V*, a state-of-the-art word embedding method which is a deep neural network trained on over a billion Google news documents [13]. Using this net, we compute a 300-dimensional vector by mapping each tag of a post to its word2vec representation. From there, average pooling is used to obtain the final 300-dimension representation of the post. At the end we use the *Textual-Sentiment* as explained in Section 2.2.

Evaluation metric: After predicting the popularity of each post at test time, we compute the Spearman's rank correlation between the prediction and ground truth which returns a value between [-1, 1]. A value close to 1 corresponds to perfect correlation.

3.1 Experiments

Experiment 1: ***Effect of multimodal features on post popularity*** In this experiment we evaluate the effect of different baseline features on predicting the popularity of a post in *Category-mix* and *Category-specific* datasets. In the *Category-specific* setting we report the average rank correlation score of all six brand categories. We also report the effect of combining *Visual features*, *Visual-Combi*, and *Textual features*, *Textual-Combi*, by average pooling in a late fusion scenario. We also report the result of combining *Visual-combi* and *Textual-combi*, denoted by *Visual-Textual-combi*.

Experiment 2: ***Effect of the engagement parameters on post popularity*** In this experiment we evaluate the post popularity prediction accuracy of our proposed *engagement parameters* on both settings of the dataset. Here, we attempt to quantify the extent to which the individual *engagement parameters* impact the popularity of a post. Finally by average pooling the rank of each post at test time we produce the *Late-fusion* results. To have a fair comparison with baseline features, we do not consider the follower parameter in the rank aggregation.

Experiment 3: ***Analysis of the engagement parameters on the popularity of each category*** In this experiment we analyze the effect of each *engagement parameter* on the popularity of brand-related posts in the *Category-specific* setting. We aim to understand which parameters make a brand become more popular. In particular, we try to understand the difference between the popularity of different brand categories. By analyzing the *engagement parameters* on brand-related posts we divide the brands into four categories: *Brand-based, Product-based, Social-based*, and *Individual-based*. The categorization is based on the parameter having the highest influence on the brand popularity, namely the presence of Brand identity (e.g., a logo), Product (e.g., fries or burger), People, and Person.

4. RESULTS

Experiment 1: ***Effect of multimodal features on post popularity*** We show the results of experiment 1 in Table 1. As we can see, the rank correlation performance using *Visual-Combi* reaches 0.233 and 0.270 in *Category-mix* and *Category-specific* respectively. In both settings, we find that the best visual feature is *Visual-Sentiment* which demonstrates the sentiment captured in the visual content of an image has a high impact on post popularity. We also observe that using *Textual-Combi* the results reach 0.231 and 0.250 on *Category-mix* and *Category-specific* respectively. To investigate whether the visual and textual features are

Table 2: Experiment 2: The result of our proposal for post popularity detection.

Datasets	*Brand*	*Product*	*Sentiment*	*Face*	*Concepts*	*Image ascetics*	*Person-Product*	*People-Product*	*Follower*	*Late-fusion without Follower*
Category-mix	0.322	0.263	0.305	0.321	0.187	0.192	0.382	0.401	0.542	0.441
Category-specific	0.361	0.293	0.335	0.375	0.196	0.203	0.414	0.425	0.555	0.462

complementary, we also combine *Visual-Combi* and *Textual-Combi* in a late fusion fashion. The result shows that they are indeed complementary, reaching 0.252 and 0.290 rank correlation accuracy on *Category-mix* and *Category-specific*.

The results of experiment 1, in general, confirm the importance of using both visual and textual features for predicting the popularity of a post in social media.

Experiment 2: ***Effect of the engagement parameters on post popularity*** We show the results of using different *engagement parameters* on the *Category-mix* and *Category-specific* in Table 2. The results demonstrate that our proposed approach has a positive effect on predicting popularity of posts where using the existence of a *Brand* results in 0.322 and 0.361 rank correlation accuracy on both *Category-mix* and *Category-specific* respectively. We also observe that by considering only the existence of a *Face* results in 0.321 and 0.375 in both settings. The results of post popularity prediction by *Person-Product* and *People-Product* in both settings show the effectiveness of these variables as compared to considering *Face* and *Product* separately. We observe that the fast food brand-related posts showing one person or people with products are more attractive and will likely get a higher number of likes. By fusing the rank correlation scores of our *engagement parameters* on each post by the average operator, the result reaches 0.441 and 0.462 popularity detection accuracy in *Category-mix* and *Category-specific* settings.

The results show a significant improvement in post popularity prediction against the baseline features. The results improve by a factor of two, 0.441 versus 0.252 on *Category-mix* and 0.462 versus 0.290 on *Category-specific*. Here we compare our proposal with *Visual-Textual-combi* which as shown in experiment 1 is the best performing baseline. It shows that for predicting the popularity of brand-related posts considering visual and textual features alone is not sufficient. The results confirm that post popularity prediction accuracy profits from using *engagement parameters*, which incorporate the correlation of different parameters with the popularity of posts.

Experiment 3: ***Analysis of the engagement parameters on the popularity of each category*** We plot the result of experiment 3 in Figure 2. The rank correlation of late fusion over all *engagement parameters* reaches to 0.520, 0.450, 0.424, 0.451, 0.421, and 0.511 for brand categories *Burger King, Culver's, Jack In The Box, Sonic Drive In, Wendy's,* and *McDonald's* respectively. By focusing on the correlation of the *Brand* parameter with the popularity of each category we find that for *McDonald's* and *Burger King*, the existence of the logo of the brand in user posts is more important in comparison with the other brands. This suggests that these two categories are more *Brand-based* in comparison with *Culver's* and *Jack In The Box* which are *Product-based*. We also observe that the existence of *Face* in all brand categories has a high impact on post popularity. It has a particularly high impact when combined with the *Product* parameter. The correlation score between the *People-Product* parameter and the popularity of *Burger King, Culver's, Jack In The Box, Wendy's*, and *McDonald's* shows the value of the social and cultural aspects of the brands. The posts for these brands show that users tend to enthusiastically socialize while consuming their product. On the other hand, the comparison of the rank correlation score of *Person-Product* and *People -Product* on *Sonic Drive In* shows the user's individual relationship with the product.

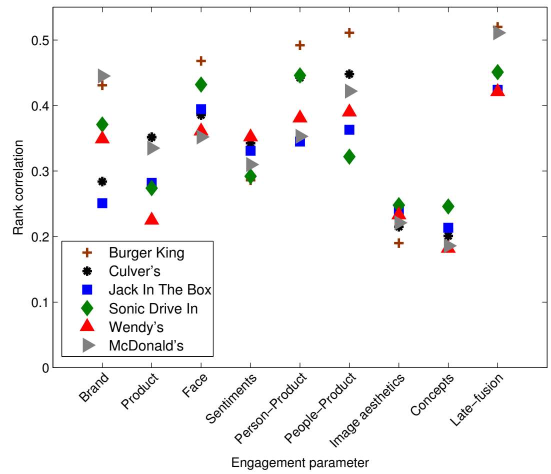

Figure 2: Experiment 3: The result of *engagement parameters* in prediction of brand post popularity for all six categories.

The results show that our approach is able to differentiate between the popularity of brands by analyzing the effect of *engagement parameters* on brand related user posts.

5. CONCLUSION

In this paper we study the drivers behind popularity of brand-related social media posts. Different from existing work, which simply models the popularity of a post directly from the visual and text features, in our approach we extract a set of parameters which are known to influence post popularity based on research frameworks in the filed of Marketing and directly embed them into the post popularity prediction model. Our proposed *engagement parameters* capture a number of factors of critical importance for studying the popularity of multimodal posts in a marketing setting. We evaluate the performance of our proposal by conducting a set of experiments on a collection of brand-related Instagram posts. The results of our experiments show that popularity prediction using the *engagement parameters* outperforms direct modelling of popularity using visual and text features only by up to a factor of two. Our *engagement parameters*, such as the existence of brand logo, sentiment and image aesthetics, all play important role in predicting post popularity. Moreover, the results show that utilizing the user *engagement parameters* gives us the ability not only to better predict the popularity of a post, but also to highlight the properties specific of each brand category.

Acknowledgments This research is supported by the Amsterdam Academic Alliance Data Science (AAA-DS) Program Award to the UvA and VU Universities.

6. REFERENCES

[1] Y. Bae and H. Lee. Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers. *J. Am. Soc. Inf. Sci. Technol.*, 63(12):2521–2535, 2012.

[2] D. Borth, R. Ji, T. Chen, T. Breuel, and S.-F. Chang. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *MM*, 2013.

[3] S. Cappallo, T. Mensink, and C. Snoek. Latent factors of visual popularity prediction. In *ICMR*, 2015.

[4] L. de Vries, S. Gensler, and P. S. H. Leeflang. Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, 26(2), 2012.

[5] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[6] F. Gelli, T. Uricchio, M. Bertini, A. Del Bimbo, and S.-F. Chang. Image popularity prediction in social media using sentiment and context features. In *MM*, 2015.

[7] K. Ginsberg. Instabranding: Shaping the personalities of the top food brands on instagram. *EJURC*, 6(1):78–91, 2015.

[8] L. Hong, O. Dan, and B. D. Davison. Predicting popular messages in twitter. In *WWW*, 2011.

[9] A. Khosla, A. Das Sarma, and R. Hamid. What makes an image popular? In *WWW*, 2014.

[10] M. Mazloom, A. Habibian, D. Liu, C. G. M. Snoek, and S.-F. Chang. Encoding concept prototypes for video event detection and summarization. In *ICMR*, 2015.

[11] M. Mazloom, X. Li, and C. G. M. Snoek. Tagbook: A semantic video representation without supervision for event detection. *TMM*, 18(7):1378–1388, 2016.

[12] P. J. McParlane, Y. Moshfeghi, and J. M. Jose. Nobody comes here anymore, it's too crowded; predicting image popularity on flickr. In *ICMR*, 2014.

[13] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *NIPS*. 2013.

[14] A. Smith, E. Fischer, and C. Yongjian. How does brand-related user-generated content differ across youtube, facebook, and twitter? *Journal of Interactive Marketing*, 26(2), 2012.

[15] G. Szabo and B. A. Huberman. Predicting the popularity of online content. *Commun. ACM*, 53(8):80–88, 2010.

[16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *CVPR*, 2015.

[17] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment in short strength detection informal text. *J. Am. Soc. Inf. Sci. Technol.*, 61(12), 2010.