A Spatio-Temporal Category Representation for Brand Popularity Prediction

Gijs Overgoor
Amsterdam Business School
University of Amsterdam
g.overgoor@uva.nl

Masoud Mazloom
Informatics Institute
University of Amsterdam
m.mazloom@uva.nl

Marcel Worring
Informatics Institute
University of Amsterdam
m.worring@uva.nl

Robert Rietveld
Amsterdam Business School
University of Amsterdam
r.l.rietveld@uva.nl

Willemijn van Dolen
Amsterdam Business School
University of Amsterdam
w.m.vandolen@uva.nl

ABSTRACT

Social media has become an important tool in marketing for companies to communicate with their consumers. Firms post content and consumers express their appreciation for the brand by following them on social media and/or by liking the firm generated content. Understanding the consumers' attitudes towards a particular brand on social media (i.e. liking) is important. In this paper, we focus on a method for brand popularity prediction and use it to analyze social media posts generated by various brands during a specific period of time. Existing instance-based popularity prediction methods focus on popularity of images, text, and individual posts. We propose a new category based popularity prediction method by incorporating the spatio-temporal dimension in the representation. In particular, we focus on brands as a specific category. We study the behavior of our method by performing four experiments on a collection of brand posts crawled from Instagram with 150,000 posts related to 430 active brands. Our experiments establish that 1) we are able to accurately predict the popularity of posts generated by brands, 2) we can use this post-level trained model to predict the popularity of a brand, 3) by constructing category representations we are

ACM Reference format:
Gijs Overgoor, Masoud Mazloom, Marcel Worring, Robert Rietveld, and Willemijn van Dolen. 2017. A Spatio-Temporal Category Representation for Brand Popularity Prediction. In *Proceedings of ICMR '17, June 6–9, 2017, Bucharest, Romania, ,* 9 pages.
DOI: http://dx.doi.org/10.1145/3078971.3078998

1 INTRODUCTION

Social media has become an integral part of people's lives and understanding the drivers behind popularity of content on social media brings us closer to understanding people's interests, opinions, and general behavior. Therefore, in recent years the prediction of the popularity of online content has been studied extensively

ICMR '17, June 6–9, 2017, Bucharest, Romania
© 2017 ACM. ISBN 978-1-4503-4701-3/17/06…$15.00
DOI: http://dx.doi.org/10.1145/3078971.3078998

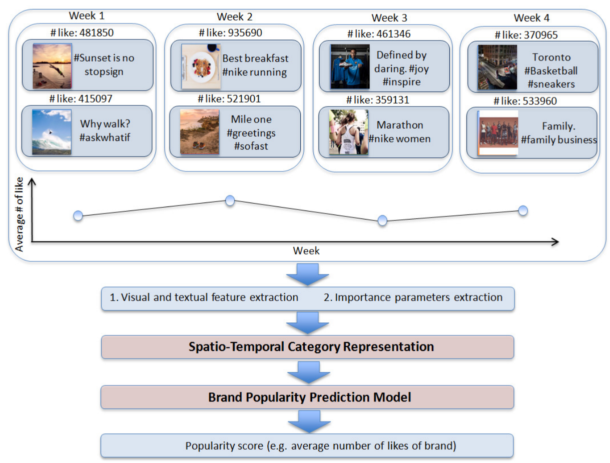

Figure 1: We propose a spatio-temporal category representation to model the temporal behavior of category specific content and use it for predicting the popularity of the category.

[1, 3, 6, 7, 9, 12, 13]. However, modeling the user's interests is difficult as they are partly expressed implicitly via the visual and the textual content of the post the user generates. A further complicating factor is that the interests of the users change over time. Because of this [8], one would expect the popularity of content posted by brands to change over time as well. And indeed, some factors of popularity such as user activeness and image prevalence variability can be viewed as time-sensitive [20], which means that for a most accurate analysis of popularity one needs to address temporal dynamics. Various automated systems have been developed to extract important information from different types of user generated content using domain agnostic methods. These automated systems have given us the ability to gain insight in what is posted by users on a large scale. True insight, however, is only possible when it has a direct impact within the domain of study [15].

One particular domain in which insight in social media has an enormous impact is business and marketing. Social media enables businesses to reach thousands of users at the same time and opens a line of communication between them and consumers which is far more interactive and direct than what was available before. Analyzing and effectively using social media holds numerous opportunities for enhancing customer engagement and brand development. Remarkable research efforts show that social media has dramatically altered the way companies market their businesses. In [10], Mangold *et al.* state that social media as a new channel and context, is crucial in today's marketing-mix. Zeng *et al.* in [22] describe how businesses profit from using social media as a source of information for business intelligence as well as an execution platform for product design and innovation, relationship management, and marketing. In [16, 17], the authors show the importance of analyzing social media from a business perspective. In [16], Risius *et al.* demonstrate the value of social media analytics for building brand loyalty. All these works [10, 16, 17, 22] use several statistics of user posts for analyzing different problems in business without analyzing their contents.

One of the interesting and challenging problems in the marketing and business community now is predicting the popularity of a brand based on the content of what they post. Brand popularity reflects a brand's recognition on social media, which comes from positive experiences of consumers. Understanding the drivers behind the popularity, therefore enables brands to enhance social media strategies for establishing a better relationship with consumers. Recent research on predicting the popularity of a user post has focused on visual and textual content [1, 3, 6, 7, 9, 12, 13], with methods depending on the type of platform and social context. Elaborating on the success of recent post popularity prediction in social media [1, 3, 6, 7, 9, 12, 13] we continue the study by predicting the popularity of brands. We follow [12] for analyzing the content generated by brands. In this paper, for the study of brand popularity prediction, we use the content of posts generated by the brands themselves instead of user posts for representing brands.

Up to now, focal priority has been cross-sectional modeling of popularity based on users and the content they generate, however as indicated before temporal dynamics are also of importance as users' interests change over time [8, 20]. Motivated by [8, 20], we propose a novel approach for incorporating the temporal dimension of content posted by brands to predict their popularity. In this paper we, therefore, incorporate the temporal dimension in the brand representation to account for changes over time in popularity that are not solely reflected by the content. We attempt to answer to the question: *How do the spatial and temporal properties of the content generated by a brand affect its popularity?*. Figure 1 visualizes our proposal for representing a brand and how we utilize it for the brand popularity prediction problem.

Of course popularity depends on how users express their opinions and this varies over the different social media platforms. Among those platforms, Instagram is used for self-expression by images with a description through captions and hashtags. Unlike some other social media platforms, the content that is generated by the brands will always be visible for users. As a consequence Instagram allows for visual marketing that is not controlled. Moreover, the hashtags allow brands to target individuals and engage with them directly. Consequently, top brands on Instagram are seeing a per-follower engagement[1] (i.e. likes per follower) rate of 4.21, which is 58 times higher than Facebook and 120 times higher than Twitter. These aspects make Instagram of particular interest in the research of brand popularity prediction. In this paper we, therefore, investigate brand popularity prediction on a dataset crawled from Instagram.

We make the following main contributions in this paper:

- We show the effectiveness of post popularity prediction on content generated exclusively by brands.
- We initiate the study of brand popularity prediction based on brand generated content on social media.
- We propose a spatio-temporal category representation for brand popularity prediction.
- We introduce a new dataset, for brand popularity prediction, obtained for free from Instagram by a simple crawling procedure.

We organize the remainder of this paper as follows. We start by considering related work in Section 2. Section 3 describes our problem formulation and we define our brand representation for predicting the popularity of brand. We introduce the experimental setup on our dataset in Section 4. Results are presented in Section 5. Finally, Section 6 concludes with a summary of our findings and a discussion of several possible directions for future work.

2 RELATED WORK

In the past few years, extensive research has been done in analyzing the content generated by users on social media. In this section, we first review existing studies in post popularity prediction, then we explain those works that focus on post popularity prediction using temporal information.

2.1 Post Popularity Prediction

Popularity prediction of user generated posts in social media has recently received a lot of attention from the research community [1, 3, 6, 7, 9, 12, 13]. While some of the work has focused on predicting the popularity of textual content, such as messages or tweets on Twitter [1, 7], recent research focuses on the image content [3, 6, 9, 12, 13].

Notable examples of text based popularity prediction are [7] and [1]. In [7], Hong *et al.* predict the number of retweets on twitter using textual features extracted from tweets. They report the effect of combining textual features with contextual features of the user as well as temporal dynamics of retweet chains and suggest the importance of including a temporal dimension into popularity prediction. Bae *et al.* in [1] report the relation between the sentiment and popularity of tweets. These works successfully predict popularity of posts using textual data. However, visual content, which is rich in information, is not addressed.

In [3, 6, 9, 12, 13], the authors focus primarily on extracting different visual features from posts for predicting the popularity of an image and measuring the impact of these features on popularity. In [9], Khosla *et al.* report the results of image popularity prediction using simple image features, such as color, gist, gradient, texture, and low-level and high-level deep learning features as indicators for objects in images. They consider the problem of popularity prediction as a learning to rank problem. In [13], McParlane *et al.* learn a binary classifier for investigating the popularity of a post. They report the effect of social factors of each post such as

[1]http://blogs.forrester.com/nate elliott/14 04 29 Instagram is the king of social engagement

how many followers a user has, the number of tags attached to the photo, and the length of the title. Moreover, they report the result of popularity prediction using content factors such as the number of faces in the images, analysis of the scene, and color features. Cappallo *et al.* in [3] learn a ranker by considering popular and unpopular latent factors. In [6, 12] the authors investigate the effect of sentiment analysis on the popularity of a post. Gelli *et al.* in [6] investigate the effect of visual sentiment analysis on images as well as contextual features used in [13]. They report the potential of predicting the popularity of images using visual sentiment scores as a feature, which was first introduced in [2] to detect sentiment in an image. Mazloom *et al.* in [12] initiate the problem of predicting popularity of brand-related user posts automatically in the business and marketing community. They further propose usage of an ensemble of cues, extracted from visual and textual channel of posts, which are important in analyzing brand popularity. All these works [6, 9, 12, 13] emphasize that visual and textual features are complementary in predicting the popularity of user generated posts. However, these works [3, 6, 9, 12, 13] don't address the use of temporal dynamics in popularity prediction. Different from these works, we propose to incorporate the temporal dimension of brand generated posts since multiple factors of popularity are time sensitive [20]. Moreover, [3, 6, 9, 12, 13] are limited to popularity prediction of individual posts. We aim for popularity of categories based on multiple posts.

2.2 Temporal Post Popularity Prediction

In [21], Yang *et al.* analyze the temporal dynamics in social media content. They propose the k-spectral centroid algorithm for clustering time series to find patterns in social media. McParlane *et al.* in [13] utilize time, day and season in their representation of posts for predicting popularity. The time at which content is posted can be classified into time of day (i.e. morning, afternoon, evening, night), whether it is a week or a weekend day and also the season in which the content is posted. Wu *et al.* in [20] argue how time plays a crucial role in social media popularity. To capture the temporal dynamics of image popularity, they factorize popularity in the user-item context and the time-sensitive context. Different from these works [13, 20, 21] which focus on considering temporal dynamics on user generated post popularity prediction, we aim to incorporate the temporal dimension of posts generated by a brand to construct a category representation for predicting the popularity of a brand.

3 OUR PROPOSAL

In this paper, we aim to predict popularity of a category by introducing a category representation that incorporates temporal dynamics. To that end we consider each individual brand as a category throughout this work. Our category popularity prediction framework consists of two main parts, schematically illustrated in Figure 2. In the training phase, at first, we construct a category representation from a dataset of content generated by categories. We use the average number of likes of categories as labels. Second, in the prediction phase, we construct a representation of new categories and compute a popularity score. Before introducing the

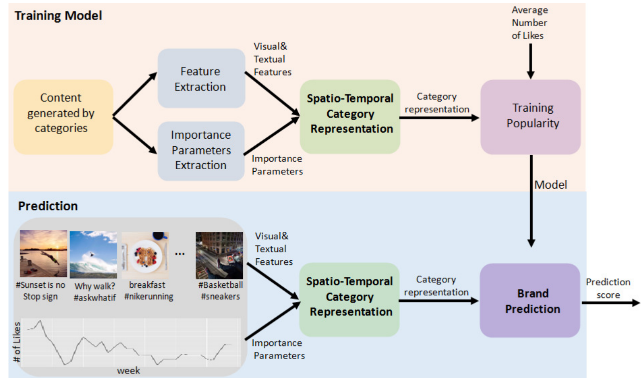

Figure 2: Dataflow for learning a category representation and using it for brand popularity prediction.

proposed framework for popularity prediction, the notation and key concepts will be formally introduced.

3.1 Problem Formalization

Given a specified category, popularity prediction of a category is the task of computing a score that shows how popular the category will be in comparison to other categories, based on the content that they generate. For consistency, we use B_x to indicate the given category x. We aim to construct a real-valued function $f(B_x)$ which produces a score for the popularity of B_x. By sorting all categories in a test set according to $f()$ in descending order, a list of most popular categories will be obtained.

Let $B = \{(B_1, y_1), (B_2, y_2), ..., (B_m, y_m)\}$ be a set of m categories, where y_i is the popularity score of category B_i. Suppose category B_i has been shared in total n_i posts during the specific period of time; $B_i = \{P_{i1}, P_{i2}, ..., P_{in_i}\}$, where P_{i1} and P_{in_i} are the first and last post shared by the category. Each post $P_{ij}, j = 1, ..., n_i$ has received a certain number of likes, α_{ij}, during a specific time which defines the popularity of post P_{ij}. Let $l_i = \{\alpha_{i1}, \alpha_{i2}, ..., \alpha_{in_i}\}$ be the set of popularity scores of B_i posts during the specific time period. We define $y_i = f(B_i) = \frac{1}{n_i}\sum_{j=1}^{n_i} \alpha_{ij}$, the average number of likes of the category's posts, as the popularity score of B_i.

The difficulty in constructing $f(B_x)$ largely depends on the representation of category B_x. We hypothesize that what makes a category become popular on the web, captured in the average number of likes a category will receive, depends on the content the category generates. The key idea of our proposal is to represent a category, based on the content shared on the web, as accurate as possible before predicting it's popularity.

Next, we show in section 3.2 how to construct a spatio-temporal category representation based on content generated and posted by a category. For the ease of reference, Table 1 lists the main notation used throughout this work.

3.2 Category Representation

Let us now explain how a specific category B_i can be represented using its generated content during a specific time. We propose to construct a category representation, $h(B_i)$,by a set of posts $P_{ij}, j = 1, ..., n_i$ generated and shared by B_i, α_{ij} the relative importance of P_{ij} based on the number of likes users gave to P_{ij}, and $\gamma_{it}, t =$

Table 1: Main notations used in this work

Notation	Definition
B	a set of categories
B_i	a set of posts of the i^{th} category
P_{ij}	j^{th} post of category B_i
n_i	number of posts shared by category B_i
α_{ij}	number of likes which P_{ij} received
γ_{ij}	importance of category B_i in j^{th} week
y_i	popularity score of category B_i
$h(B_i)$	a representation of category B_i
$f()$	a function for computing the popularity

$1, ..., K$ the relative importance of B_i per week based on the moving average popularity time series, with $\sum_{t=1}^{K} \gamma_{it} = 1$ for all categories. Suppose $P_{ij} = (P_{ij}^v, P_{ij}^t)$ consists of two components namely visual and textual content respectively. To find different representations of post P_{ij}, we extract visual features, $z(P_{ij}^v)$, and textual features, $z(P_{ij}^t)$, of P_{ij} and use them for representing B_i.

Similar to [12], we use both visual and textual components in the representation of category B_i as they carry some information related to the popularity of a category. In addition, we make use of a third component, a temporal dimension, with which we aim to capture changing interests in time of users. Depending on how to use these components, we present three variants of category representation $h(B_i)$, namely

(1) Naive Category Representation (NCR): The construction of the naive brand category representation is simply an average pooling over all features of posts of B_i without considering importance parameters. The representation is defined as follows:

$$h_{NCR}^v(B_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} z(P_{ij}^v) \tag{1}$$

where n_i is the total number of posts generated by category B_i.

(2) Weighted Category Representation (WCR): The construction of the weighted category representation is based on the weekly importance parameter γ_{it}. The weekly importance parameter is derived from the popularity time series of a category and it indicates the relative importance of a category per week compared to the other weeks. The representation is defined as follows:

$$h_{WCR}^v(B_i) = \sum_{t=1}^{K} \gamma_{it} \sum_{j \in I_{it}} \frac{z(P_{ij}^v)}{n_{it}} \tag{2}$$

where I_{it} is the set of posts of B_i within week t. K is the total number of weeks in the time series and n_{it} is the total number of posts generated by B_i in week t. The second sum in the formula shows the average pooling of visual features of posts per week.

(3) Spatio-Temporal Category Representation (STCR): This representation is based on both importance parameters, γ_{it}, and α_{ij}. The weighted aggregation based on the importance parameters is given by the following equation:

$$h_{STCR}^v(B_i) = \sum_{t=1}^{K} \gamma_{it} \sum_{j \in I_{it}} \frac{\alpha_{ij} * z(P_{ij}^v)}{A_{it}} \tag{3}$$

where A_{it} is the total number of likes B_i received in week t.

Replacing $z(P_{ij}^v)$ by $z(P_{ij}^t)$ in the above three equations, we produce $h_{NCR}^t(B_i)$, $h_{WCR}^t(B_i)$, and $h_{STCR}^t(B_i)$, three category representations using textual components of posts of B_i.

Next, we show in section 3.3 how to tackle brand popularity prediction using the different category representation methods.

3.3 Brand Popularity Prediction

After representing brands with the category representation, we use it for the problem of brand popularity prediction. Suppose $B = B_{TR} \cup B_{TE}$ is a set of m brands with $B_{TR} = \{(B_1, y_i), (B_2, y_2), ..., (B_e, y_e)\}$ is a training set consist of e brands and $B_{TE} = \{(B_{e+1}, y_{e+1}), ..., (B_m, y_m)\}$ is a test set consisting of the other brands in B. By dividing the brands in a train and test set using the formulas explained in section 3.2, we define B_{TR} and B_{TE} as two matrix representations of all brands:

$$\begin{aligned} B_{TR} &= [h(B_1), h(B_2), ..., h(B_e)] \\ B_{TE} &= [h(B_{e+1}), h(B_{e+2}), ..., h(B_m)] \end{aligned} \tag{4}$$

Each row of B_{TR} and B_{TE} represents a brand . We train a brand popularity model on B_{TR} and report the result of popularity prediction on B_{TE}.

Let $h(B_i)$ be the category representation of brand B_i and y_i show the popularity of B_i. The idea is to optimize w, parameter of function $f_w()$, on B_{TR} to minimize the error between y_i and $f_w(h(B_i)) = w^T h(B_i)$. We consider the problem of brand popularity prediction as a regression problem and try to optimize the following objective function:

$$\sum_{i=1}^{e} (y_i - f_w(h(B_i))) - C \sum_{k=1}^{n} {w_k}^2 \tag{5}$$

which can be formulated as

$$\arg\max_w \sum_{i=1}^{e} \log p(y_i | h(B_i), w) - C \sum_{k=1}^{n} {w_k}^2 \tag{6}$$

where $\log p(y_i | h(B_i), w) = \frac{1}{1+e^{-w^T h(B_i)}}$.

To solve the problem and find the optimal value of w we use L2 regularized L2 loss Support Vector Regression, as used in [9, 12], from the LIBLINEAR package [5]. After training the model and finding the optimum value of w on B_{TR}, we use it for prediction of brand popularity on B_{TE} and report the rank correlation between the predicted scores and grand truth.

4 EXPERIMENTAL SETUP

We investigate the effectiveness of our proposal for predicting popularity of a brand by performing a series of experiments on a dataset crawled from Instagram.

Figure 3: Example of three brands, *Nike, Ford, Expedia* and their generated content during four weeks from our dataset.

4.1 Dataset

Since there is no existing dataset for predicting the popularity of a brand, we created one by crawling the content brands post on Instagram. A grand variety of brands have adopted Instagram for the means of visual marketing. Unlike other social media platforms, the content that is generated by the brands will always be visible by users. Moreover, the hashtags allow brands to target certain individuals and engage with them directly. This makes our Instagram dataset especially relevant for the research on brand popularity prediction. Our dataset consists of 150,000 posts generated by 430 active brands, such as *Expedia, Ford, Nike* and *McDonald's*, across 27 different industries like *Travel, Automotive, Sports* and *Fast Food*. All the content is generated between 01/05/2015 and 30/04/2016 and the brands behind the content are considered active, which means that they at least generate content every week. Figure 3 shows examples of posts with the number of likes generated by three highly popular brands. [2]

The dataset is split into two parts, where two thirds are used for training the model and one third is used as a test set for evaluation of the model.

4.2 Implementation details

Feature extraction To represent a post in our dataset we use the following state of the art features:

Textual features We use the textual features used in [12]:

- *W2V* A trained deep neural network proposed in [14], which computes a 300 dimension vector by mapping each tag onto its Word2Vec representation. A post is then represented by average pooling of the W2V representation of all tags in the post.
- *Textual Sentiment* Represented by making use of SentiStrength [19]. First the stop words are removed and stemming is performed. The sentiment value of each tag is computed using SentiStrength to generate positive (ranging from 1 to 5) and negative (ranging from -1 to -5) sentiment scores.
- *Terms* a sparse representation with a length of 5250, where the length of the vector equals the number of unique tags that occur in at least 75 posts. Only the values of those elements corresponding to the tags of the post receive a weighting score.

Visual features We use the visual features as used in [12]:

- *CNN-Pool5* As low-level visual features, we used the 1024-dimensional features from pooling the last fully connected layer of the Deep Net in [18] which is trained on ImageNet [4]
- *Concepts* we extract the convolutional neural network features, initially proposed in [18] and trained to identify those 15,293 ImageNet [4] concept categories, for which at least 200 positive examples are available. We represent the image of each post by the 15,293-dimensional output of the softmax layer of the network as an effective representation has been shown in [11].
- *Visual Sentiment* A feature based on the Visual Sentiment Ontology which was first introduced in [2] to detect sentiment in an image. The sentiment ontology representation

[2] http://isis-data.science.uva.nl/Masoud/ICMR17Data

consist of a 1,200 dimensional vector with probabilities of Adjective Noun Pairs (ANP) being present in the image.

Popularity prediction As we mentioned in section 3.3, we use an L2 regularized L2 loss Support Vector Regression to predict the popularity of a brand. To find the optimal value of the regularization parameter C, we consider 5-fold cross-validation on C for $C \in \{0.001, 0.1, 1, 10, 100\}$.

Evaluation metric After predicting the popularity of each brand at test time we compute the Spearman's rank correlation between the prediction and ground truth, the average number of likes of the posts which the brand shared in Instagram, which returns a value between [-1, 1]. A value close to 1 corresponds to perfect correlation.

4.3 Experiments

Experiment 1: ***Post popularity prediction*** In this experiment we evaluate the popularity of the brand generated posts. We train a post level popularity model over all posts in the training set and use it for predicting the popularity of a post in the test set. We report the result of post popularity prediction using different visual and textual features, which we mentioned in section 4.2, on our test set. We also report the result of fusing visual and textual features, by an average pooling scenario, for predicting the popularity of brand generated post. The combination of features creates an overview of the current state of the art prediction methods on content generated exclusively by brands.

Experiment 2: ***Baseline: Brand popularity by post level training model*** We create and evaluate a baseline for brand popularity prediction using a post level training model. We use this model to predict the popularity score for each post of a brand in the test set. We obtain a popularity score for a brand in the test set by averaging the predicted score of each post of a brand. Then we report the rank correlation between the predicted scores and ground truth as a popularity of a brand. We also report the result of the brand popularity prediction using different visual and textual features and their fusion by average pooling.

Experiment 3: ***Brand popularity prediction using brand category representation*** In this experiment we evaluate the three category representation methods we explained in section 3.2 for predicting the popularity of a brand. We represent each brand in the train and test set using our proposal which is explained in section 3.3 and train a model on the train set and predict the popularity of a brand in the test set. We evaluate the effect of different visual and textual features for representing a brand category in brand popularity prediction. We also report the result of fusing all features by average pooling.

Experiment 4: ***Popularity prediction on an off-line collection*** In this experiment we evaluate the effect of our proposal for selecting those images from an off-line collection of images of a brand which are expected to get a high number of likes. For this purpose we randomly select 100 posts of each brand from the test set and use only the visual data. We extract the visual features per test image and apply the pre-trained models of experiment 1 to compute a popularity score. Then, we rank the images of the brand from the test set based on their popularity score and select the top p images as most promising for sharing in social networks. We evaluate the selection, by defining the popularity ratio as average number of matches between images selected by our method and images from the ground truth ranking. The value of the popularity ratio shows the quality of the approach for selecting images to be shared in social networks. Proximity of the popularity ratio to 1 indicates a better image selection. We report the accuracy of selecting popular images using different visual features by selecting 10, 20, 30, ..., 100 images for sharing in social networks. We also repeat this procedure by selection images of brands randomly. We repeat the random selection of images for 50 times and report the average of the results.

5 RESULTS

5.1 Post Popularity Prediction

We report the result of this experiment in Table 2. Starting with the textual features, the results show the 0.346, 0.481 and 0.145 rank correlation as a popularity of a brand generated post using *W2V, Terms,* and *Textual Sentiment* features respectively. Using an average pooling for fusing these textual features, the result reaches 0.494. The results in Table 1 also show the importance of visual features in predicting the popularity of a post where the results reach 0.341, 0.287, 0.293 and 0.366 rank correlation using *CNN-Pool5, Concepts, Visual Sentiment* and *fusion* of them respectively. The results depict that *Concepts* and *Sentiment* present in the image of a post holds information on the popularity of a post. People may respond to concepts in images positively or negatively and a post that contains those positive concepts generally attains more likes. At the end the result in Table 1 shows the significant improvement in predicting the popularity of a brand generated post by combining the textual and visual features. The result of rank correlation reaches 0.520.

The results of experiment 1 confirm that visual and textual features are complementary for predicting the popularity of a brand generated post. Moreover, in general, we are able to predict popularity of brand generated content.

5.2 Baseline: Brand popularity by post level training model

Table 3 shows the results of Experiment 2, where we use a post-level training model on predicting popularity of a brand. It shows the rank correlation reaches 0.474, 0.133, 0.143, and 0.479 using *W2V, Terms, Textual Sentiment*, and *fusing* these textual features respectively. The Table 3 also mention to the effect of visual feature on brand popularity, where the rank correlation reaches 0.408, 0.299, 0.399 and 0.411 using *CNN-Pool5, Concept, Visual Sentiment*, and *fusing* of visual features respectively. All visual features have a significant positive impact on brand popularity. Table 3 also confirms complementarity of the visual and textual features for brand popularity prediction by training a post-level popularity model, where the result reaches 0.500 rank correlation.

The results of experiment 2 show the ability of predicting the popularity of a brand using a trained model based on the content they generated. Again, the results confirm that visual and textual features are complementary for predicting the popularity of a brand.

Table 2: The result of predicting the popularity of brand generated posts.

	Textual Features				Visual Features				Multimodal
Features	W2V	Terms	Sentiment	Fusion	CNN-Pool5	Concept	Sentiment	Fusion	**Fusion**
Rank-correlation	0.346	0.481	0.145	0.494	0.341	0.287	0.293	0.366	**0.520**

Table 3: Brand popularity prediction by post level training model.

	Textual Features				Visual Features				Multimodal
Features	W2V	Terms	Sentiment	Fusion	CNN-Pool5	Concept	Sentiment	Fusion	**Fusion**
Rank-correlation	0.474	0.133	0.143	0.479	0.408	0.299	0.399	0.411	**0.500**

Table 4: Category representation for brand popularity prediction.

	Textual Features				Visual Features				Multimodal
Features	W2V	Terms	Sentiment	Fusion	CNN-Pool5	Concept	Sentiment	Fusion	**Fusion**
NCR	0.517	0.141	0.154	0.527	0.441	0.322	0.445	0.479	**0.554**
WCR	0.526	0.167	0.174	0.532	0.464	0.365*	0.461	0.497	**0.579**
STCR	0.532	0.197	0.204	0.540	0.487	0.394	0.471	0.511	**0.598**

5.3 Brand Popularity prediction by category representation

We show the result of experiment 3 in Table 4. Table 4 describes the result of brand popularity using our proposal for brand representation using visual, textual, and fusing of them.

Using *NCR* as a brand representation the result reaches 0.517, 0.141, 0.154 and 0.527 when we use *W2V, Terms, Textual Sentiment*, and *fusing* them respectively. On the other hand using visual features for the *NCR* representation of a brand, the result of brand popularity reaches 0.441, 0.322, 0.445 and 0.479 when we use *CNN-Pool5, Concepts, Visual Sentiment*, and *fusion* of all visual features respectively. Combining visual and textual features in the *NCR* representation of a brand, the rank correlation reaches 0.544.

The results in Table 4 also show the efficiency of using *WCR* as a method for brand representation in comparison with *NCR* method. We observe that the rank correlation using all textual and visual features reaches 0.526, 0.167, 0.174, 0.464, 0.365, 0.461 where features are *W2V, Terms, Textual Sentiment, CNN-Pool5, Concepts*, and *Visual Sentiment* respectively. By fusing textual features, visual features, and combining visual and textual features the rank correlation reaches to 0.532, 0.497, and 0.579 respectively. We observe 5% relative improvement in brand popularity prediction compared to *NCR*, showing the effect of considering the importance per week inside the brand representation.

The results in Table 4 emphasize the efficiency of considering the temporal behaviour of popularity of contents generated by a brand inside the brand representation. Using *STCR*, the rank correlation reaches 0.532, 0.197, 0.204, 0.540 using *W2V, Terms, Textual Sentiment*, and *fusing* them respectively. By visual feature to construct *STCR* the rank correlation reaches 0.487, 0.394, 0.471, 0.511 where the features are *CNN-Pool5, Concepts, Visual Sentiment*, and *fusion* respectively. The Table 4 show the best result of brand popularity reaches 0.598 where we fused the visual and textual features represented by *STCR*. By comparing the result of *STCR* by *NCR* and *WCR* we find that incorporating the importance of brand per week and importance of post inside the weak, we reach to more accurate brand representation.

The STCR addresses not only the weekly importance of certain images but also assigns weights to images within the week. It emphasises specifically the images of importance for the brand and combines it with the importance of a certain week in the brand popularity over time.

By comparing the result of our category representations with the baseline, experiment 2, we find 10%, 16%, and 20% relative improvement in brand popularity prediction when we use *NCR, WCR*, and *STCR* respectively.

The results of experiment 3 confirm that brand popularity prediction accuracy profits from constructing a representation of a brand. Moreover it depicts that a brand representation in which temporal behavior is incorporated is more representative for popularity prediction.

5.4 Popularity prediction on off-line collection

We show the results of experiment 4 in Figure 4. The results demonstrate the effectiveness of using different visual features for selecting those images from an offline collection which have the potential of getting more likes and becoming more popular. Figure 4 shows the result of image selection using our proposal always much better than random selection. When we request to select 30 images we reach 0.125 accuracy in popularity ratio using *Random* selection, whilst using *Concepts, Sentiments, CNN-Pool5* and *Fusion* of all visual features reach 0.280, 0.328, 0.371, and 0.410. The result shows that by considering fusion of all visual features, we come to an accurate set of selected images which depicts each type of visual features capture different aspects of popularity of a post.

We also observe from Figure 4 that the presence of *Visual Sentiment* for all brands has a positive effect on the predicted popularity. We further investigate which particular ANPs correlate most with the number of likes in different industries. In order to evaluate the correlation of visual sentiment with popularity, we compute the SVR weights separately for two industries, *Food* and *Fashion*, and

Figure 5: Experiment 4: ANPs that have a positive impact on popularity of (a) *Food industry* and (b) *Fashion industry.* Font size correlates with the importantnees of sentiments for both brands.

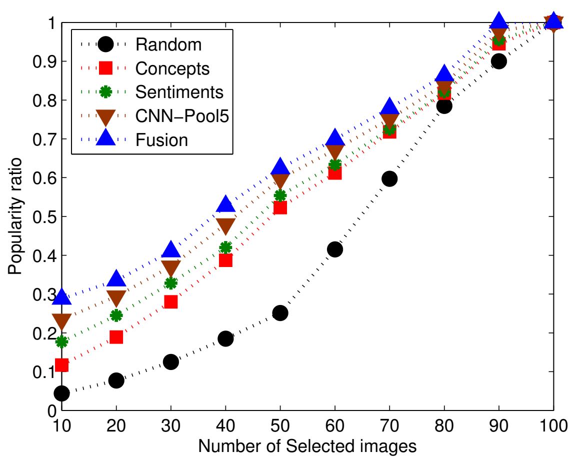

Figure 4: Experiment 4: Selection of images from an offline collection of brands predicted as more popular. Fusing all visual features help us to select more popular images.

then sort them for selection. The results in Figure 5, highlight the important sentiments with different impact on popularity for these brand industries.

The results of experiment 4 show that, in general, using a predicted model over the brand generated contents we have the ability of helping brands in selecting the content they should share on their page on social media.

6 CONCLUSION

In this paper we study the prediction of popularity of brands. We investigate how the visual and textual contents of brand generated posts impact their popularity. Different from existing works which focus more on predicting the popularity of user generated posts, we are the first to study brand popularity prediction. We incorporate the spatio-temporal behavior of the popularity of the posts generated by brand for representing brand. We study the behavior of our proposal for predicting the popularity of brands on a dataset crawled from Instagram.

The results of experiment 1 confirm complementarity of visual and textual features for predicting post popularity. In addition, it shows how we are able to accurately predict the popularity of content that is generated and posted exclusively by brands. The results of experiment 2 show that by using the post level training model we have the ability to predict the popularity of the brand. Experiment 3 displays that prediction of brand popularity is more accurate when a brand is represented as a category representation. Moreover, incorporation of the temporal dimension into the representation increases the predictability of brand popularity. Finally, experiment 4 reveals that using our proposal, where all channels are fused, we have the ability of selecting a set of off-line images of a brand likely to become more popular.

We conclude that for category popularity prediction it is beneficial to construct a category representation in which spatio-temporal dynamics are considered.

ACKNOWLEDGMENTS

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

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. G. Snoek. Latent factors of visual popularity prediction. In *ICMR*, 2015.
[4] 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.
[5] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *JMLR*, 9:1871–1874, 2008.
[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] L. Hong, O. Dan, and B. D. Davison. Predicting popular messages in twitter. In *WWW*, 2011.
[8] B. Jiang and Y. Sha. Modeling temporal dynamics of user interests in online social networks. *Procedia Computer Science*, 51:503–512, 2015.
[9] A. Khosla, A. Das Sarma, and R. Hamid. What makes an image popular? In *WWW*, 2014.
[10] W. G. Mangold and D. J. Faulds. Social media: The new hybrid element of the promotion mix. *Business horizons*, 52(4):357–365, 2009.
[11] 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.
[12] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. Van Dolen. Multimodal popularity prediction of brand-related social media posts. In *MM*, 2016.
[13] 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.

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

[15] C. North. Toward measuring visualization insight. *IEEE Comput. Graph. Appl.*, 26(3), 2006.

[16] M. Risius and R. Beck. Effectiveness of corporate social media activities in increasing relational outcomes. *Information Management*, 52(7):824–839, 2015.

[17] B. Stieglitz, Dang-Xuan and Neuberger. Social media analytics - an interdisciplinary approach and its implications for information systems. *Business Horizons*, 6(2):89–96, 2014.

[18] 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.

[19] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment strength detection in short informal text. *J. Am. Soc. Inf. Sci. Technol.*, 61(12):2544–2558, 2010.

[20] B. Wu, T. Mei, W.-H. Cheng, and Y. Zhang. Unfolding temporal dynamics: Predicting social media popularity using multi-scale temporal decomposition. In *AAAI*, 2016.

[21] J. Yang and J. Leskovec. Patterns of temporal variation in online media. In *WSDM*, 2011.

[22] D. Zeng, H. Chen, R. Lusch, and S.-H. Li. Social media analytics and intelligence. *Intelligent Systems*, 25(6):13–16, 2010.