Towards Tweet Content Suggestions for Museum Media Managers

Marco Furini
University of Modena and Reggio Emilia
Dip. Comunicazione ed Economia
Reggio Emilia, Italy
marco.furini@unimore.it

Federica Mandreoli
University of Modena and Reggio Emilia
Dip. Scienze Fisiche, Informatiche e Matematiche
Modena, Italy
federica.mandreoli@unimore.it

Riccardo Martoglia
University of Modena and Reggio Emilia
Dip. Scienze Fisiche, Informatiche e Matematiche
Modena, Italy
riccardo.martoglia@unimore.it

Manuela Montangero
University of Modena and Reggio Emilia
Dip. Scienze Fisiche, Informatiche e Matematiche
Modena, Italy
manuela.montangero@unimore.it

ABSTRACT

Cultural Heritage institutions are embracing social technologies in the attempt to provide an effective communication towards citizens. Although it seems easy to reach millions of people with a simple message posted on social media platforms, media managers know that practice is different from theory. Millions of posts are competing every day to get visibility in terms of likes and retweets. The way text, images, hashtags and links are combined together is critical for the visibility of a post. In this paper, we propose to exploit machine learning techniques in order to predict whether a tweet will likely be appreciated by Twitter users or not. Through an experimental assessment, we show that it is possible to provide insights about the tweet features that will likely influence its reception/recommendation among readers. The preliminary tests, performed on a real-world dataset of 19,527 museum tweets, show promising accuracy results.

CCS CONCEPTS

• **Human-centered computing → Social media**; *Empirical studies in collaborative and social computing*;

KEYWORDS

Twitter, machine learning, prediction

ACM Reference Format:
Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2018. Towards Tweet Content Suggestions for Museum Media Managers. In *International Conference on Smart Objects and Technologies for Social Good (Goodtechs '18), November 28–30, 2018, Bologna, Italy*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3284869.3284874

Goodtechs '18, November 28–30, 2018, Bologna, Italy

ACM ISBN 978-1-4503-6581-9/18/11.
https://doi.org/10.1145/3284869.3284874

1 INTRODUCTION

In the last few years, social technologies have changed our personal and professional life: they entered and changed almost every aspect of our society, from health to entertainment, from work to leisure, from education to business [3]. People use these technologies for many different reasons: to socialize, to post personal opinions about products and services, to improve their visibility, to share their thoughts and experiences. Although criticized for privacy issues [15], with no doubts, the importance of social media in people's life has been increasingly recognized and new ways of exploitation are rapidly emerging for different purposes [6, 13, 14, 20, 21].

Cultural Heritage (CH) is among the sectors that might receive great benefits from social technologies [8, 19]. Indeed, cultural operators and organizations have the opportunity to advertise their initiatives in an easy and simple way and Twitter is being used by many CH institutions [7, 22].

However, if on the one side it is easy to use these platforms, on the other side it is difficult to get noticed in a ocean of messages. That is, it is not easy to write a tweet to motivate people to visit a museum. [9, 12]. To clarify, although a message is mainly composed of a text, hashtags, and links, it is worth noting that a text might be subjected to different linguistic structures (i.e., the same semantic content might be written in different ways), a hashtag might be composed in different ways, and a link might connect the message to external resources like images and video. In this scenario, the main challenge that a media manager has to transform a semantic content into a successful tweet (i.e., high number of retweets and likes).

Motivated by the need to support effective communication and marketing campaigns within the CH domain, our goal is to design and develop an innovative dashboard that will guarantee a continuously-updated analysis of the Twitter account and that will provide media managers with effective content *suggestions* for writing successful tweets. For instance, the manager might receive suggestions like: "Use #visitVanGogh together with #VanGogh: visibility will likely increase of 30%", or "Use mentions at the end of the message rather than at the beginning: 45% of readers will likely retweet it".

In this paper, we propose a predictive strategy leveraging on tweet content that will be a piece of the engine of the innovative dashboard. Through a content-based predictive approach, we aim to: i) foresee the popularity of the tweet the media manager is writing, and ii) give insights on the content features that most influenced the prediction (e.g., "use of the word *conversation* has not brought benefits 7 times out of 10").

In particular, we introduce a predictive approach that exploits well-known machine learning techniques for automatic message classification and performs an exploratory analysis on a real-world dataset (19,527 museum tweets) to experimentally verify the feasibility of the idea. It is worth noting that, in this preliminary study, prediction is performed in a simplified scenario where two classes are defined: GOOD if the message will likely be appreciated by readers, BAD otherwise.

The rest of the paper is organized as follows: after an overview of the state of the art (Section 2), Section 3 describes how we gathered and selected the data and the features and how we prepared the dataset. The experiments on predictive analysis are presented in Section 4, while Section 5 concludes the paper.

2 RELATED WORK

Recently, different studies focused on the use of social media in the CH sector. For instance, [5] proposes a quantitative and qualitative approach to the analysis of tweets posted during the MuseumWeek event organized by Twitter; [4] introduces a set of KPIs for quantitative estimation of CH sensitivity as expressed by social network users; in [22] eleven Twitter performance indexes are used to describe the activity and performance of the top-60 European museums and their Twitter accounts; [18] seeks to understand more about the relationship building that museums are engaging in using Twitter by measuring a set of content and frequency parameters of a sample of U.S. museums on Twitter; [10, 11] analyze tweet contents extracted from museum accounts to investigate what features (e.g., images, hashtags, mentions, links, etc.) are worth using. The studies considered popular tweets posted by official museum accounts or ordinary people and analyzed them to derive insights about the tweet generation.

In line with the papers above, this paper focuses on the cultural sector and, in particular, on art museums. However, the goal of this work is different. It is not a coarse grain analysis of the use of social media but rather a fine grain analysis of the features that characterize tweets of museum accounts and their use in predicting tweet influence.

Different approaches have been proposed in the literature to analyze and predict tweet influence, a.k.a. popularity. In most cases, prediction deals with machine learning approaches. For instance, paper [24] focused on news agencies accounts on Twitter and studies the propagation characteristics of news on Twitter as a backbone of a Twitter news popularity prediction model. In this study, they also found that the negative sentiment of news has some correlation with tweet popularity while the positive sentiment does not have such obvious correlation. The work in [25] concerned with a popular micro-blogging website in China, Sina Weibo, and aims to discover content factors and contextual factors that affect the popularity of tweets. They found that the two factors are equally important to predict the first popularity measure, i.e. the tweets diffusion, but content outperforms context when predicting the second popularity measure, i.e., the number of comments tweets received. In [16], authors aimed to identify features for tweet popularity prediction that are both effective and effortless, i.e., easy to obtain or compute. From the experimental assessment, it followed that a relative small set of features, in particular temporal features, can achieve comparable performance to all features.

While all these papers focused on a notion of tweet popularity that is influenced by the network, our aim, instead, is mainly to study content features and their impact on popularity prediction. Moreover, the focus of our work is on art museums that, as to our knowledge, have never been studied in this context.

The impact of multimedia content on tweet popularity and life span was studied in [26]. The study showed that multimedia tweets dominate pure text both because they are more popular and because they survive longer. Finally, sentiment analysis in Twitter is a field that has recently attracted research interest. An overview of the main algorithms that have been proposed for sentiment analysis in Twitter is provided in [17], whereas [23] investigated whether the community sentiment energy of a topic is related to the spreading popularity of the topic. Experiments on two communities found a linear correlation between the community sentiment energy and the real spreading popularity of topics. In our work [11], instead, we noticed that tweets originated by museums are typically neutral so in this paper we do not further investigate on this feature.

3 DATASET AND FEATURE DEFINITION

Museum Selection and grouping. We selected 25 well known world spread art museums and clustered them according to the number of followers of their Twitter account. We got six groups that are listed in Table 1 together with their number of followers (at the time we were observing data). Note that, in the clustering process, we considered the number of followers as this number reflects the potential visibility that tweets written by the museum account might have. Therefore, to have comparable situations we divided the accounts in such groups. To clarify, a tweet that receives 10 likes might be considered as attractive if wrote by a single person, but the judgment is different if the tweet was written by a museum with millions of followers. Therefore, when analyzing tweets, we concentrate only on those features that are proper of the tweet and do not depend on the context around the sender. In numbers, we analyzed around 800 tweets authored by each of the 25 museums, for a total of 19,527 tweets.

Tweet feature selection. In this paper, we do not consider those features that are related to the account originating the tweet (e.g., the number of followers of the author account), but we use the ones that can be drawn from the tweet itself (i.e., content features). Therefore, among the commonly used tweet features (see e.g. [2, 25]), we empirically selected those that resulted more relevant in the museum environment: we performed a preliminary experimental evaluation where we investigated different features (length, sentiment, hashtags, URLs, etc.) and discarded those that were transparent to the classification (i.e., the classification did not change if taken or

Table 1: Museum groups

Group 1 (G1)	# Followers	**Group 2 (G2)**	# Followers	**Group 3 (G3)**	# Followers
@MuseumModernArt	5,120,000	@britishmuseum	1,560,000	@CentrePompidou	970,000
@Tate	4,500,000	@vangoghmuseum	1,330,000	@NationalGallery	887,000
@metmuseum	3,680,000	@MuseeLouvre	1,250,000	@museofrodakahlo	847,000
@Guggenheim	3,350,000	@GettyMuseum	1,250,000	@MuseeOrsay	610,000
@saatchi_gallery	2,800,000	@museodelprado	1,180,000		
Group 4 (G4)	# Followers	**Group 5 (G5)**	# Followers	**Group 6 (G6)**	# Followers
@mfaboston	332,000	@visitmuve_it	88,000	@MuseoEgizio	21,600
@museiincomune	264,000	@museupicasso	65,000	@Uffizi	19,900
@philamuseum	247,000	@mart_museum	64,000	@MUSE_Trento	12,600
@maspmuseum	245,000				
@ngadc	216,000				
@Museo_MAXXI	190,000				

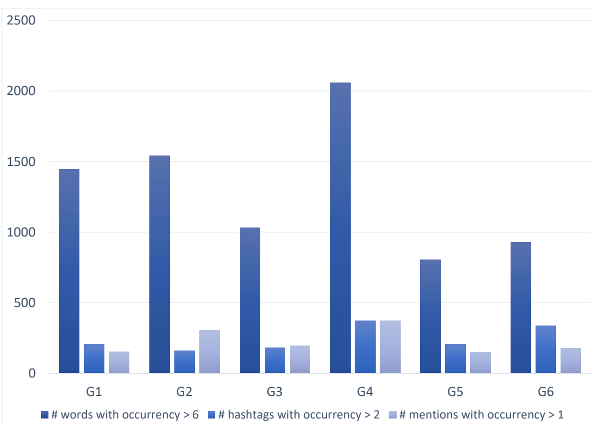

Figure 1: For each group, number of terms (words, hastags and mentions) of the vocabulary appearing with a given frequency in the set of tweets.

not taken into consideration). We eventually selected the following content features:

- Countable: the number of *hashtags* (i.e., words preceded by #) in a tweet, the number of *URLs* (i.e., links to external resources), the number of *media* contents (i.e., image, video, graphical emoticons, . . .), the number of *mentions* (i.e., twitter account preceded by). Table 2 reports statistics of such numbers within the six museum groups.
- Frequent vocabulary: the terms (words, hashtags and mentions) that appeared with a certain frequency in the full text of the tweet. In particular, we consider the words that appear more than six times in the full tweet corpora, the hashtags that appear more than twice and the mentions that appear more than once. Figure 1 reports the cardinalities of such sets for each museum group.

Finally, we considered the source of the tweet (i.e., the museum account that generated the tweet).

Dataset preparation. One of the aims of this study is to understand whether the proposed approach is worth pursuing and, thus, if it actually works. Therefore, we start with a simple binary tweet classification into GOOD and BAD ones: we determine whether a tweet is good or bad by means of the number of likes it got, group by group. Intuitively, if the number of likes of a given tweet of a museum in a given group is close to the maximum number of likes received by the tweets of the same group, then the tweet is classified as GOOD, and dually for BAD ones. Moreover, to avoid ambiguity in tweet classification, we mainly work with a dataset that does not contain those tweets for which such classification is hard even for humans. For example, let's say that in a given group the maximum (resp., minimum) number of likes a tweet received is ℓ_{max} (resp. ℓ_{min}), then it is natural to say that a tweet that received ℓ_{max} minus a small constant likes has been appreciated by users, but what would humans say of a tweet that received a number of likes close to $(\ell_{max} + \ell_{min})/2$? Would we say it has been appreciated or not? Therefore, for each group we select the 20% of the tweets that received the highest number of likes to be the set of GOOD tweets, and analogously, the 20% with smallest number of likes to be the BAD ones. In conclusion, in the dataset there is the same number of GOOD tweets and BAD tweets. Moreover, we have that, a GOOD tweet has more than 219, respectively 300, 87, 38, 16 and 51, likes for a museum belonging group G1, respectively G2, G3, G4, G5 and G6; a BAD a tweet has less than 56, respectively 52, 20, 8, 3 and 3, likes for a museum belonging to group G1, respectively G2, G3, G4, G5 and G6. At the end of the following section, we will discuss what happens when considering a "dirtier" dataset.

4 PREDICTIVE ANALYSIS

The system uses machine learning techniques for automatic message classification. To verify the feasibility of the system, we performed several experiments in order to evaluate the impact of the various features, including content features vs source information, different modeling/interpretations of the numerical features and the impact of tweet vocabulary. The reference classifier exploited for all tests is the Naïve Bayes one; the GOOD and BAD classes are defined with a threshold of 20% of the total messages as discussed in Section 3. Furthermore, we evaluated the impact of considering

Table 2: Tweet Features. `Min` **is not reported as it is equal to zero for all groups and each feature.**

	# URLs	# Hashtags	# media c.	# mentions	# URLs	# Hashtags	# media c.	# mentions
	G1				**G2**			
Average	0,678	0,996	1,169	0,353	0,631	1,032	1,483	0,512
Max	3	9	10	7	3	14	14	9
	G3				**G3**			
Average	0,793	1,223	1,340	0,536	0,606	1,356	1,410	0,531
Max	3	12	7	10	3	15	14	16
	G5				**G6**			
Average	0,613	1,963	1,330	0,859	0,428	2,091	1,100	0,694
Max	3	11	37	14	4	18	8	16

larger parts of the original datasets by extending the GOOD and BAD classes beyond the 20% threshold. In a final test we will also analyze the performance of using different classification algorithms (i.e., Decision Tree and Max Entropy) [1]. Training set and test set are randomly chosen from the considered set of tweets in ratios of 4:5 and 1:5, respectively, w.r.t. the selected number of tweets. All the classification accuracy figures are derived as an average of 20 runs. The prototype implementation of our system is written in Python and exploits the Natural Language Toolkit (NLTK) library[1].

The first experiment (E1) is conducted by considering our reference set of countable content features (number of media, hashtags, mentions, URLs) and the tweet source. The obtained results show that the accuracy (percentage of correct classifications) levels are quite high, especially for some of the groups. Indeed, the achieved accuracies are: 0.735 (G1), 0.893 (G2), 0.749 (G3), 0.794 (G4), 0.744 (G5) and 0.974 (G6). It is to note the high value of the accuracy achieved within G6.

To verify the impact on accuracy, we performed other experiments on different sets/interpretations of features. The second experiment (E2) is conducted with two binary interpretations of the features: (a) absence/presence of a feature; (b) feature value below/above the median. Results are reported in Figure 2. It is to note that the accuracy is not hardly affected by these changes as the accuracy values are very similar to the ones achieved through experiment E1 (just +1% for G3 and +2.5% for G5).

The third experiment (E3) is conducted with a configuration similar to that of E1, except for the source of the tweets. In this experiment, we remove the source information to understand its impact in the classifier prediction accuracy. Figures 3 shows the obtained results. It can be noted that the source helps the classifier as it introduces important benefits (+4 % for G1; +9 % for G2; +15 % for G3; +3 % for G4; +5 % for G5 and +6% for G6). Consequently, we can state that, even if the content features typically bring the most important contribution in discriminating tweets, using the source information (when possible) can bring good accuracy improvements.

The fourth experiment (E4) is conducted by considering the vocabulary among the features. In particular, we considered three different kinds of vocabularies: only hashtags, only mentions, all the words (see also Section 3). Figure 4 shows the obtained results. If on the one side, hashtags and mentions did not contribute to increase the accuracy with respect to the results obtained in E1, on the other side it is possible to note that, by considering the presence of the specific words of the tweets as features, the accuracy increases for most of the groups (+7% for G1; +2% for G3; +3% for G4 and +5%

[1]http://www.nltk.org

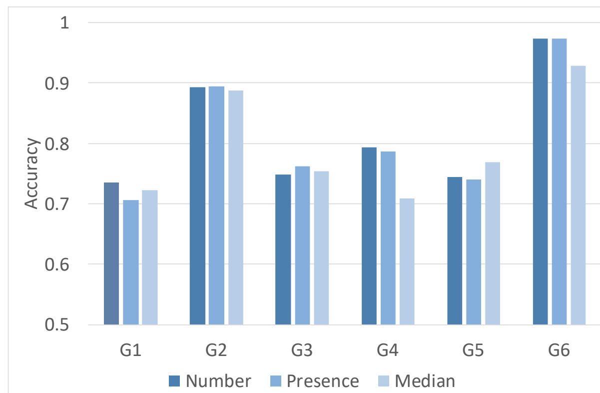

Figure 2: Accuracy comparison among different feature interpretations (E2): number, presence and above/below the median.

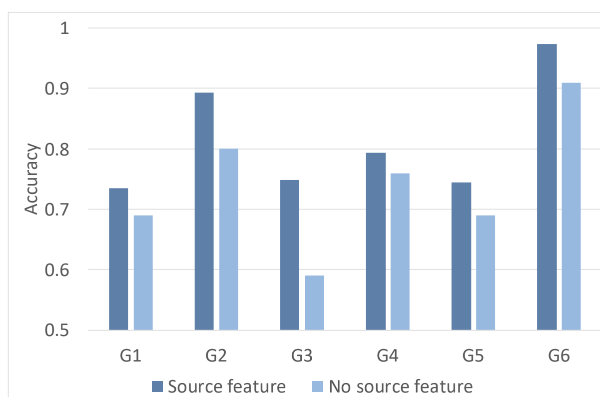

Figure 3: Accuracy comparison (E3): the influence of the source.

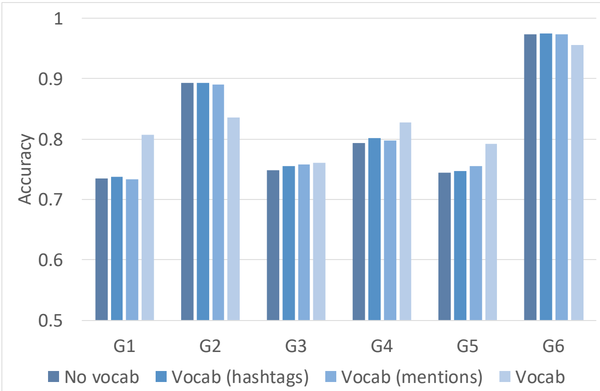

Figure 4: Accuracy comparison (E4): the influence of the vocabulary.

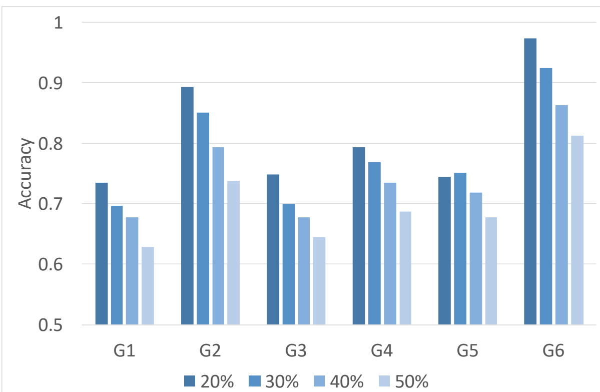

Figure 5: Accuracy comparison on different dataset compositions (E5): changing the GOOD and BAD classes size.

for G5), and it decreases for just two groups (-5% for G2 and -2% for G6).

The fifth experiment (E5) aims to investigate how accuracy varies w.r.t. different compositions of the considered dataset. In particular, we evaluated the impact of considering larger parts of the original datasets by extending the GOOD and BAD classes beyond the 20% threshold. The obtained results are shown in Figure 5. As it was to be expected, gradually including a larger and larger number of tweets that are difficult to classify leads to a slow degradation of the performance. When considering the complete dataset (i.e., GOOD and BAD thresholds set to 50%), we loose about 10-15 points percentage, even if for some of the groups the final accuracy can still be considered as quite satisfying (above 0.7 for G2 and even above 0.8 for G6).

The sixth experiment (E6) investigates how other classifiers fare w.r.t. the achieved accuracy. In particular, we considered, in addition to the Naïve Bayes, the Decision Tree and the Max Entropy classifiers. Figure 6 shows the obtained results. Results are different from group to group, but the difference is still very small. However, in some cases it is advisable to use a different classifier than the Naïve Bayes.

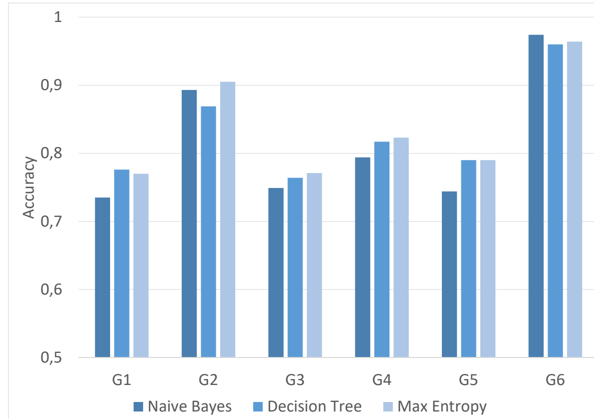

Figure 6: Accuracy comparison (E6): the performances of three different classifiers.

4.1 Discussion

The experiments we have conducted and described in this paper show that significant results can be achieved in classification accuracy and, moreover that it is possible to increase the achieved accuracy by varying the features on which the classifier bases its choice.

By analyzing the results, we can notice that content features and source can indeed drive an effective prediction of the success of a tweet. As to the specific configurations of the classifying task, there is no single configuration that fits best for all the considered groups. This is not surprising as the groups represent museums with very different characteristics.

What we foresee for a real scenario in which the museum media manager uses the system, is that (s)he will have to identify, through an automatic experimental phase, what are the features and/or the classifiers that produce a better accuracy for that particular group to which the museum belongs. Once these features are identified, the system will support the museum media manager according to the results obtained by the identified classifier and by the identified features.

Furthermore, by identifying the features that most influenced the choice of the classifier, the system can give easy-to-understand insights on the tweet features that will likely influence the acceptance of the message among readers. Table 3 shows an example of the results given for binary content features (above/below median interpretation) and their behavior group by group (see also E2 in Section 4. For example we can see that, for all groups, writing a tweet with too many (i.e., above median) URLs is definitely not advisable: for instance, for G1, bad tweets having this feature were 9.8 times more frequent than those that went good. Moreover, a museum of the G6 group should use a high number of media because, within the dataset, the G6 tweets went bad if they did not contain abundant media content.

To check if there are words inside the message that will be accepted/refused by readers, the classifier highlights the words that most contributed to the classification of tweets. This allows

Table 3: Most informative features for different groups (boolean content features, above/below median interpretation).

	URLs		Hashtags		media content		mentions	
	above med.	below med.	above med.	below med.	above med.	below med.	above med.	below med.
G1	👎 9.8:1.0	-	👎 1.2:1.0	-	👎 1.5:1.0	👍 1.1:1.0	👎 3.2:1.0	👍 1.5:1.0
G2	👎 18.3:1.0	👍 1.1:1.0	👎 2.0:1.0	👍 1.2:1.0	👎 1.1:1.0	👍 1.1:1.0	👎 6.3:1.0	👍 2.5:1.0
G3	👎 6.0:1.0	-	👎 2.0:1.0	👍 1.3:1.0	-	-	👎 1.3:1.0	👍 1.2:1.0
G4	👎 5.0:1.0	-	👍 1.2:1.0	👎 1.1:1.0	👍 2.2:1.0	👎 1.3:1.0	👎 1.3:1.0	👍 1.1:1.0
G5	👎 4.7:1.0	👍 1.1:1.0	👍 1.8:1.0	👎 2.1:1.0	👍 1.4:1.0	👎 1.1:1.0	👎 1.1:1.0	👍 1.1:1.0
G6	👎 13.7:1.0	👍 1.1:1.0	👍 3.7:1.0	👎 4.5:1.0	👍 3.0:1.0	👎 1.3:1.0	-	-

the system to build sets of words that will be accepted/refused by readers. For example, for the G1 group it is not recommended to use the word "conversation" (negative trend ten times greater). Similarly, for the G4 group, the word "ticket" should be avoided (a nine times higher negative trend).

5 CONCLUSIONS AND FUTURE WORK

In this paper, we investigated the use of machine learning to predict the success of museum tweets. In particular, the prediction is based on content features and not on the context ones.

Although preliminary, the obtained results showed that it is possible to achieve good prediction accuracy. Given such promising preliminary results, in the future we plan to: increase the prediction accuracy by considering three classes instead of two (e.g., GOOD, BAD and NEUTRAL) and by analyzing the tweets linguistic structure.

ACKNOWLEDGMENTS

This work is partially supported by the University of Modena and Reggio Emilia within the FAR 2016 Department Project "SocialGQ".

REFERENCES

[1] E. Alpaydin. 2010. *Introduction to Machine Learning*. MIT Press. https://books.google.it/books?id=4j9GAQAAIAAJ

[2] Roja Bandari, Sitaram Asur, and Bernardo A. Huberman. 2012. The Pulse of News in Social Media: Forecasting Popularity. In *Proceedings of the Sixth International Conference on Weblogs and Social Media*.

[3] Armir Bujari, Marco Furini, Federica Mandreoli, Riccardo Martoglia, Manuela Montangero, and Daniele Ronzani. 2017. Standards, Security and Business Models: Key Challenges for the IoT Scenario. *Mobile Networks and Applications* (2017), 1–8. https://doi.org/10.1007/s11036-017-0835-8

[4] A. Chianese, F. Marulli, and F. Piccialli. 2016. Cultural Heritage and Social Pulse: A Semantic Approach for CH Sensitivity Discovery in Social Media Data. In *Proc. of the IEEE 10th International Conference on Semantic Computing (ICSC)*. 459–464.

[5] A. Courtin, B. Juanals, J. L. Minel, and M. d. S. LÃĪger. 2014. The Museum Week Event: Analyzing Social Network Interactions in Cultural Fields. In *Proc. of the 10th International Conference on Signal-Image Technology and Internet-Based Systems*. 462–468.

[6] Roberta De Michele, Stefano Ferretti, and Marco Furini. 2018. On helping broadcasters to promote TV-shows through hashtags. *Multimedia Tools and Applications* (Aug 2018). https://doi.org/10.1007/s11042-018-6510-7

[7] Nicola Dusi, Ilaria Ferretti, and Marco Furini. 2016. PlayTheCityRE: A visual storytelling system that transforms recorded film memories into visual history. In *2016 IEEE Symposium on Computers and Communication (ISCC)*. 85–90. https://doi.org/10.1109/ISCC.2016.7543719

[8] Nicola Dusi, Ilaria Ferretti, and Marco Furini. 2017. A transmedia storytelling system to transform recorded film memories into visual history. *Entertainment Computing* 21 (2017), 65 – 75. https://doi.org/10.1016/j.entcom.2017.05.002

[9] Mike Ellis. 2011. *Managing and Growing a Cultural Heritage Web Presence: A Strategic Guide*. Facet Publishing.

[10] Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2017. The Use of Hashtags in the Promotion of Art Exhibitions. In *Digital Libraries and Archives - 13th Italian Research Conference on Digital Libraries*. 187–198.

[11] Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2018. 5 Steps to Make Art Museums Tweet Influentially. In *Proc. of the 3rd International Workshop on Social Sensing (SocialSens)*.

[12] Marco Furini and Manuela Montangero. 2008. The Impact of Incentive Mechanisms in Multi-Channel Mobile Music Distribution. *Multimedia Tools and Applications* 37, 3 (March 2008), 365–382.

[13] Marco Furini and Manuela Montangero. 2016. TSentiment: On gamifying Twitter sentiment analysis. In *2016 IEEE Symposium on Computers and Communication (ISCC)*. 91–96. https://doi.org/10.1109/ISCC.2016.7543720

[14] Marco Furini and Manuela Montangero. 2018. Sentiment analysis and Twitter: a game proposal. *Personal and Ubiquitous Computing* 22, 4 (01 Aug 2018), 771–785. https://doi.org/10.1007/s00779-018-1142-5

[15] Marco Furini and Valentina Tamanini. 2015. Location privacy and public metadata in social media platforms: attitudes, behaviors and opinions. *Multimedia Tools and Applications* 74, 21 (2015), 9795–9825. https://doi.org/10.1007/s11042-014-2151-7

[16] Shuai Gao, Jun Ma, and Zhumin Chen. 2014. Effective and Effortless Features for Popularity Prediction in Microblogging Network. In *Proceedings of the 23rd International Conference on World Wide Web*. 269–270.

[17] Anastasia Giachanou and Fabio Crestani. 2016. Like It or Not: A Survey of Twitter Sentiment Analysis Methods. *ACM Comput. Surv.* 49, 2 (2016), 28:1–28:41.

[18] L. A. Langa. 2014. Does Twitter Help Museums Engage with Visitors?. In *Proceedings of iConference*. 484–495.

[19] Silvia Mirri, Catia Prandi, Marco Roccetti, and Paola Salomoni. 2017. Handmade Narrations: Handling Digital Narrations on Food and Gastronomic Culture. *J. Comput. Cult. Herit.* 10, 4, Article 20 (July 2017), 17 pages. https://doi.org/10.1145/3097569

[20] Lorenzo Monti, Giovanni Delnevo, Silvia Mirri, Paola Salomoni, and Franco Callegati. 2018. Digital Invasions Within Cultural Heritage: Social Media and Crowdsourcing. In *Smart Objects and Technologies for Social Good*, Barbara Guidi, Laura Ricci, Carlos Calafate, Ombretta Gaggi, and Johann Marquez-Barja (Eds.). Springer International Publishing, Cham, 102–111.

[21] Catia Prandi, Andrea Melis, Marco Prandini, Giovanni Delnevo, Lorenzo Monti, Silvia Mirri, and Paola Salomoni. 2018. Gamifying cultural experiences across the urban environment. *Multimedia Tools and Applications* (08 2018). https://doi.org/10.1007/s11042-018-6513-4

[22] Vasiliki Vrana, Kostas Zafiropoulos, and Konstantinos Antoniadis. 2016. Top European Museums on Twitter. In *Tourism and Culture in the Age of Innovation*. Springer, 457–469.

[23] Xiang Wang, Chen Wang, Zhaoyun Ding, Min Zhu, and Jiumin Huang. 2018. Predicting the popularity of topics based on user sentiment in microblogging websites. *J. Intell. Inf. Syst.* 51, 1 (2018), 97–114.

[24] Bo Wu and Haiying Shen. 2015. Analyzing and Predicting News Popularity on Twitter. *International Journal of Information Management* 35, 6 (2015), 702–711.

[25] Lun Zhang, Tai-Quan Peng, Ya-Peng Zhang, Xiao-Hong Wang, and Jonathan J. H. Zhu. 2014. Content or context: Which matters more in information processing on microblogging sites. *Computers in Human Behavior* 31 (2014), 242–249.

[26] Xun Zhao, Feida Zhu, Weining Qian, and Aoying Zhou. 2012. Impact of Multimedia in Sina Weibo: Popularity and Life Span. In *Semantic Web and Web Science - 6th Chinese Semantic Web Symposium and 1st Chinese Web Science Conference*. 55–65.