What Multimedia Sentiment Analysis Says About City Liveability

Boonzajer Flaes, J.; Rudinac, S.; Worring, M.

DOI
10.1007/978-3-319-30671-1_74

Publication date
2016

Document Version
Final published version

Published in
Advances in Information Retrieval

License
Article 25fa Dutch Copyright Act

What Multimedia Sentiment Analysis Says About City Liveability

Joost Boonzajer Flaes[1(✉)], Stevan Rudinac[2(✉)], and Marcel Worring[2(✉)]

[1] Twitter Inc., London, UK
jflaes@twitter.com
[2] Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands
{s.rudinac,m.worring}@uva.nl

Abstract. Recent developments allow for sentiment analysis on multimodal social media content. In this paper we analyse content posted on microblogging and content-sharing platforms to estimate sentiment of the city's neighbourhoods. The results of sentiment analysis are evaluated through investigation into the existence of relationships with the indicators of city liveability, collected by the local government. Additionally, we create a set of sentiment maps that may help discover existence of possible sentiment patterns within the city. This study shows several important findings. First, utilizing multimedia data, i.e., both visual and text content leads to more reliable sentiment scores. The microblogging platform Twitter further appears more suitable for sentiment analysis than the content-sharing website Flickr. However, in case of both platforms, the computed multimodal sentiment scores show significant relationships with the indicators of city liveability.

Keywords: Multimodal sentiment analysis · Semantic concept detection · Social multimedia · City liveability

1 Introduction

Posting messages on social networks is a popular means for people to communicate and to share thoughts and feelings about their daily lives. Previous studies investigated the correlation between sentiment extracted from user-generated text and various demographics [6]. However, as technology improves, the bandwidth available for users also increases. As a result, users can share images and videos with greater ease. This led to a change in types of media being shared on these online networks. More particularly, user-generated content often consist of a combination of modalities, e.g., text, images, video and audio. As a result, more recent studies have tried to predict sentiment from visual content too [2].

A recent study on urban computing conducted by Zheng et al. underlines the potential of utilizing user-generated content for solving various challenges a modern metropolis is facing with, ranging from urban planning and transportation

J.B. Flaes—This research was performed while the first author was a student at the UvA.

N. Ferro et al. (Eds.): ECIR 2016, LNCS 9626, pp. 824–829, 2016.
DOI: 10.1007/978-3-319-30671-1_74

to public safety and security [7]. In this paper we investigate whether the sentiment analysis of spontaneously generated social multimedia can be utilized for detecting areas of the city facing such problems. More specifically, we aim at creating a sentiment map of Amsterdam that may help paint a clearer picture of the city liveability. Since visual content may contain complementary information to text, in our approach we choose to utilize them jointly. Additionally, we make use of automatically captured metadata (i.e., geotags) to analyse the messages in context of the locations where they were posted.

A direct evaluation of our results would require a user study in which the participants would be asked about their sentiment on different neighbourhoods of a city. As conducting such study would be both extremely time consuming and labour intensive, here we propose the use of open data as an indirect ground-truth. We consider a large number of demographic, economic and safety parameters comprising the liveability index of a neighbourhood and investigate their association with the automatically produced sentiment scores in different scenarios.

2 Research Methodology

Our approach consists of three steps outlined in Fig. 1 and described below.

2.1 Data Collection

The data used in this research comes from two different online social networks, which both include textual and visual content. The emphasis of the analysis will be on modalities dominant on a particular platform, i.e., textual data in case of Twitter and visual data for Flickr. However, visual data shared on Twitter and text shared on Flickr will also be used for a better understanding of the content hosted on these platforms and an increased quality of the sentiment analysis.

For about two months, 64 thousand tweets were collected within a 10-mile radius of the city center of Amsterdam. The dataset only includes tweets that have a geo-location available. A total of 64 thousand images were downloaded from Flickr that are taken in and around the city of Amsterdam.

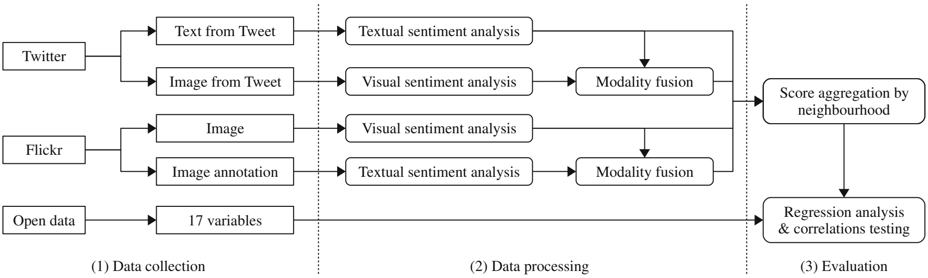

Fig. 1. The approach overview.

From open data as provided by the city of Amsterdam we utilize the following neighbourhood variables: *percentages of non-western immigrants*, *western immigrants*, *autochthonous inhabitants*, *income*, *children in households with minimum incomes*, *people working*, *people living on welfare*, *people with low, average and high level of education*, *recreation area size*, *housing prices*, *physical index*, *nuisance index*, *social index*, and *liveability index* [5].

2.2 Calculating Sentiment

Textual Data. Sentiment for the textual data is calculated by using two different Python packages. The first is NLTK which makes use of a naive Bayes classifier to predict sentiment [1]. For each tweet, the sentiment is calculated ranging from -1 to $+1$. The second is Pattern [3], which uses stemming and part-of-speech tagging to predict sentiment and includes both a Dutch and English based lexicon.

In our approach, first the language is detected using the Google Translate API. Then, if the language is not English the text is translated into English. However, for the Pattern package we also included the Dutch lexicon if a tweet was predicted to be Dutch. To compare the two packages for sentiment analysis, we manually annotated a random sample of 150 tweets as either positive or negative. On these tweets the Pattern package significantly outperformed NLTK, which is why we decided to use it for further evaluation.

Visual Content. We analyse visual content of the images using SentiBank [2] and detect 1200 adjective noun pairs (ANPs). For example, using SentiBank a 'happy person' can be detected, which combines the adjective 'happy' with the noun 'person'. Each of the ANPs detected has a sentiment score associated with it and for each image, we compute the average of the top 10 detected ANPs.

Combined Score. To obtain more reliable scores, we combine the results of visual and textual sentiment analysis. For Twitter this means that a combined (i.e., average) score will be calculated if the tweet contains a direct link to an image. For Flickr, the sentiment scores of the images are combined with sentiment extracted from the annotation text. The output of the visual sentiment classifier ranges from -2 to $+2$, whereas the output ranges from -1 to $+1$ in case of the textual sentiment analysis. We use a zero mean unit variance normalization to calculate the normalised scores.

2.3 Statistical Analysis

To find out which neighbourhood variables are related with the sentiment scores, we conducted linear regression analysis. For that, we aggregate the sentiment scores by calculating the mean score of each neighbourhood.

To identify which neighbourhood variables were significantly associated with the sentiment scores, single regression analyses were conducted for each of the neighbourhood variables and the sentiment scores. Regression coefficients were assessed for significance with statistical significance set at $p \leq 0.05$.

3 Experimental Results

We conduct several experiments in order to evaluate the proposed research methodology described in Sect. 2.3.

3.1 Flickr Sentiment Analysis

Our analysis shows that no significant relationship can be found between the visual sentiment scores from Flickr and the selected liveability indicators. The most significant relationships are found with the *percentage of people living on governmental welfare checks* and the *level of education*.

The combined sentiment scores showed more promising results. The *safety index* and the *people living on governmental welfare* showed significant associations with the sentiment scores (p = 0.037 and p = 0.028, cf. Fig. 2).

3.2 Twitter Sentiment Analysis

To compute reliable scores, only neighbourhoods with more than 40 tweets were taken into account. However, no significant relationship is found between the open data and the scores based on the analysis of textual content.

However, the combined score shows multiple significant relationships with the liveability indicators. The first interesting relationship is found between ethnic composition of the neighbourhood and sentiment scores. More particularly, there is a positive association between sentiment scores and the percentage of native Dutch inhabitants (cf. Fig. 3). Similarly, a positive association can be found between level of education and the sentiment scores. This is not surprising as these two variables are strongly correlated.

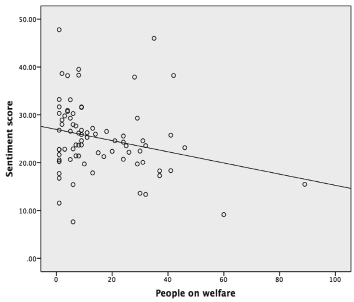

Fig. 2. Relationship between Flickr sentiment scores and percentage of people living on welfare checks.

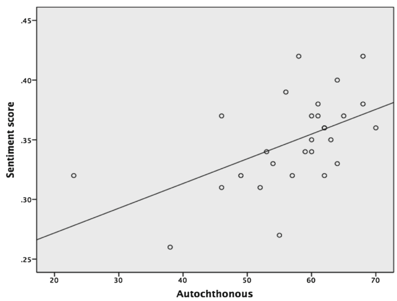

Fig. 3. Relationship between Twitter sentiment scores and percentage of autochthonous population.

3.3 Sentiment Maps

To facilitate easier gaining of new insights about the developments in the neighbourhoods and in order to visualize findings of our study, we created a simple data exploration interface[1]. The interface features interactive maps showing the city of Amsterdam and visualizing the sentiment scores of each neighbourhood generated according to the different methods evaluated in Sects. 3.1 and 3.2 (cf. Fig. 4). Additionally, the maps are visualising various indicators of city liveability for easier assessment of possible relationships (cf. Fig. 5). Finally, the interface also includes the possibility to view a sub-sample of the tweets uploaded in Amsterdam.

4 Discussion and Conclusion

In this paper we investigate if the sentiment scores derived from the analysis of social multimedia data relate to the geographic location in which they are posted. We use state-of-the-art sentiment analysis methods in order to process multimodal content from two very different social media platforms, a content-sharing website Flickr and a microblogging platform Twitter.

Our research reveals significant relationships between automatically extracted sentiment and the indicators of city liveability. Namely, both sentiment scores from Flickr and Twitter showed significant relationships with the open data when multimodal content is analysed. However, in case of both analysed platforms, we found no significant relationships when using a single modality for sentiment analysis. This confirms our assumption that the multimodal sentiment analysis provides for higher accuracy.

The detected sentiment mostly correlates with the demographics of the inhabitants. The percentage of people living on welfare checks shows a negative linear association with sentiment scores from the Flickr data. Ethnic background, income and education level show significant relationships with the sentiment

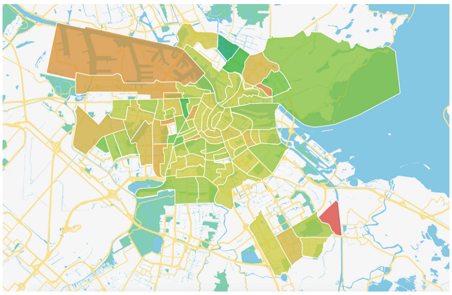

Fig. 4. Aggregated multimodal sentiment scores based on Flickr data.

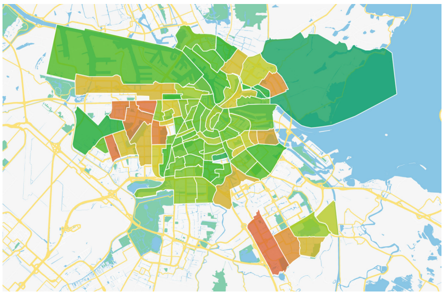

Fig. 5. Percentage of autochthonous inhabitants in Amsterdam neighbourhoods.

[1] http://goo.gl/DAj9y2.

scores based on Twitter data. Further research is needed to investigate the nature of these relationships. However, it is interesting to observe that for both platforms the economic indicators (i.e., people living on welfare checks and income) show significant relationships with our computed sentiment scores. On the other hand, the liveability or social index of a neighbourhood showed no significant relationship. Since these indices are designed to measure the subjective feelings of the inhabitants, we would have expected these to be more significant in our research.

Using the Twitter data shows more significant relationships than using the data from Flickr. A possible explanation for this might be that people do not tend to share opinions or feelings on this platform but mainly use it as a method to share their photographs.

To further improve this research, it would also be interesting to see if the sentiment prediction could be adjusted to factors that are important for residents of a city. The examples are the detectors created specifically for urban phenomena like noise nuisance or graffiti, known to influence the liveability of the city [4]. Finally, combining sentiment scores from user-generated data and open data allows for new research opportunities.

Our research shows that sentiment scores may give additional insights in a geographic area. The big advantage of training on social multimedia data is that it provides for real-time insights. Additionally, sentiment in these areas can prove to be an indication of important factors like crime rate or infrastructure quality. This may be useful for government services to know what area to improve or for new businesses to find a convenient location.

References

1. Bird, S.: Nltk: the natural language toolkit. In: Proceedings of the COLING/ACL on Interactive Presentation Sessions, pp. 69–72. Association for Computational Linguistics (2006)
2. Borth, D., Ji, R., Chen, T., Breuel, T., Chang, S.-F.: Large-scale visual sentiment ontology and detectors using adjective noun pairs. In: Proceedings of the 21st ACM International Conference on Multimedia, pp. 223–232. ACM (2013)
3. De Smedt, T., Daelemans, W.: Pattern for python. J. Mach. Learn. Res. **13**, 2063–2067 (2012)
4. Dignum, K., Jansen, J., Sloot, J.: Growth and decline - demography as a driving force. PLAN Amsterdam **17**(5), 25–27 (2011)
5. C.G.A.O. for Research and Statistics. Fact sheet leefbaarheidsindex periode 2010–2013, Febraury 2014. https://www.amsterdam.nl/publish/pages/502037/fact_sheet_6_leefbaarheidsindex_2010_-_2013_opgemaakt_def.pdf
6. Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., Danforth, C.M.: The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. PLoS ONE **8**(5), e64417 (2013)
7. Zheng, Y., Capra, L., Wolfson, O., Yang, H.: Urban computing: concepts, methodologies, and applications. ACM Trans. Intell. Syst. Technol. (TIST) **5**(3), 38 (2014)