Sentiment analysis for group decisions

TOBIAS BERGTHALER, SE Seminar in Pervasive Computing, Austria

Nowadays, it is impossible to imagine a life without social media allowing people to express themselves freely and participate in broad discussions about various topics. We want to evaluate how much agreement there is for certain statements. Furthermore, if a positive statement leads to higher agreement or not. In order to evaluate these questions, sentiment analysis is used to analyse the statements and the comments.

Additional Key Words and Phrases: group decision making, social networks, news comments, sentiment analysis

ACM Reference Format:
Tobias Bergthaler. 2021. Sentiment analysis for group decisions. *J. ACM* 0, 0, Article 0 (January 2021), 6 pages.

1 INTRODUCTION

Nowadays, social media is used for a lot of discussions. Yet it is not easy to keep track or read all appearing statements and comments. In this paper we will focus on analysing comments below a statement. We will evaluate if a certain statement gets higher agreement or disagreement in the community. For this we will use a provided dataset from SFU Opinion and Comments Corpus [Kolhatkar et al. 2018] [Kolhatkar et al. 2020]. For the decision making process we will analyse the reactions on articles provided in the dataset, therefore, we will use sentiment analysis. The reactions will be categorised in positive, negative and neutral comments. Based on this reactions we will evaluate how much (dis-) agreement a statement has. We will experiment if a positive statement receives more or less positive reactions in the comment section. In addition, we will calculate for how long statements receive comments.

This paper is structured in two main parts. Firstly, in each chapter there is a theoretical part, discussing common methods tasks and ideas. Secondly, there is a practical implementation part in some of our chapters, where our experiences while implementing will be shared.

The rest of the paper is structured as follows. In Section 2 our used literature is introduced. Section 3 introduces sentiment analysis, their methods 3.1 and our implementation of it 3.2. The idea of an agreement rating is introduced in Section 4. Finally our conclusion is shown in Section 5.

2 RELATED WORK

There exist a lot of papers and articles about sentiment analysis, in this section we introduce some of the literature we used for our research and our implementation.

The basic information for our paper we received from the following articles. The paper "Emotion Predictor Using Social Media Text and Graphology" [Roy et al. 2019] provides an overview over multiple approaches for sentiment and text analysis. The article also describes its own experiment with sentiment analysis and graphology analysis. For our article we will focus on the sentiment analysis part and implementation. An overview of

Author's address: Tobias Bergthaler, k11803976@students.jku.at, SE Seminar in Pervasive Computing, Altenberger Strasse 69, Linz, Upperaustria, Austria, 4040.

0004-5411/2021/01-ART0 $15.00
https://doi.org/

possible Approaches we could gain from the paper "Approaches, Tools and Applications for Sentiment Analysis Implementation" [D'Andrea et al. 2015] that provides an overview over different tools for sentiment analysis and usable tools. One concrete example is given in the paper "Contextual semantics for sentiment analysis of twitter" [Saif et al. 2016] that describes the problem of unknown words on twitter, for example shortcuts like "afk", "lol" etc. Furthermore, the article describes the problem of word based sentiment analysis, where words can get a different meaning in context. The paper "A lexicon model for deep sentiment analysis and opinion mining applications" [Maks and Vossen 2012] describes certain problems with simple sentiment analysis and tries a different approach. They do not just determine if a sentence is negative or positive, but assign the positive or negative meaning to a certain actor as well. Finally, an overview of the challenges is given in "A survey on sentiment analysis challenges" [Hussein 2018] that focuses on many papers and implementations and their main problems in analysing.

Our implementation is based on the article "Sentiment Strength Detection in Short Informal Text" [Thelwall et al. 2010] it describes an implementation called Sentistrength. This implementation considers spelling mistakes, grammatical incorrectness as well as shortcuts. Sentistrength is one of the best sentiment analysis tools [Saif et al. 2016]. Therefore we are going to use this implementation approach for our paper. We use the dataset [Kolhatkar et al. 2018] provided from "The SFU Opinion and Comments Corpus: A Corpus for the Analysis of Online News Comments" [Kolhatkar et al. 2020]. It consists of 10,339 articles and 663,173 comments below the articles. The information has been collected between 2012 and 2016. Unfortunately it has proven to be difficult to get a dataset that includes both, the statement (the articles in this case) and the corresponding comments. Therefore we will use this provided dataset under the Creative Commons Attribution-Non-Commercial licence.

For our decision process we used the article "Carrying out consensual Group Decision Making processes under social networks using sentiment analysis over comparative expressions" [Morente-Molinera et al. 2019] which follows the approach that experts should discuss topics on social media instead of a specialised tool. The authors[Morente-Molinera et al. 2019] follow the social media approach, because in their opinion a specialised tool has the drawback that the users need to get used to it, the user size is fixed in the beginning and each tool is different to use. Therefore, the paper suggests using social media and analysing the content between the experts of a certain topic using sentiment analysis procedures.

3 SENTIMENT ANALYSIS

There are two main approaches in sentiment analysis.

On the one hand the lexicon-based approach, where words are in lexical resources and each word is with a positive or negative value[D'Andrea et al. 2015]. Typical lexicon approaches are Sentiwords, SentiWordNet and SenticNet [Roy et al. 2019]. The lexicon-based approach works fast and does not need any further training. Furthermore, it and can be adapted fast by extending the lexical resources to a new topic.

On the other hand the machine learning approach[D'Andrea et al. 2015], were a dataset is used to train a machine learning algorithm to predict the opinion of a certain text (positive/negative). The main advantage is that this approach can adapt easily to specific topics. However, it is not easy to learn the new topics because the algorithm must be trained first. For this, the training set needs a lot of work and often human interaction (e.g.labelling the data).

Of course also combinations of lexicon-based and machine learning approaches exist. In this paper one of them is used.

3.1 Method

Sentiment analysis has different methods to analyse a text [D'Andrea et al. 2015]. The first is to analyse the whole text or document and come up with a positive or negative value. The next more detailed method is the aspect level where each aspect is analysed for it self. It is possible to go even further into detail to analyse single words and their positive or negative value. Each of the approaches has their benefits and downsides [Hussein 2018]. The word analysis has for example the problem that words get different meanings in the given context and the system does not differentiate that. So, in the sentences "It is a very good phone indeed!" and "I will leave you for good this time!" the word "good" gets the same sentiment value [Saif et al. 2016]. Aspect and whole text based methods often have their issues when it comes to grammatical and spelling mistakes [Schouten and Frasincar 2016].

In our implementation we will use Sentistrength [Thelwall et al. 2010] as the sentiment analysis tool. This tool uses a lexicon-based approach that is optimised by an machine learning algorithm. Furthermore, the algorithm can work with basic non-standard spellings and common textual writing [Thelwall et al. 2010]. Also the algorithm does not just split sentences into positive and negative sentences. There will be a range from 0 to 5 for positive and an range from -5 to 0 for negative sentences. Moreover, each sentence can contain positive as well as negative aspects. A sentence value is the sum of the word values as shown in Figure 1 received from the Sentistrength Website [Sentistrength 2020]

The text 'I love you but hate the current political climate.'
has positive strength **3** and negative strength **-4**

Approximate classification rationale: I love[3] you but hate[-4] the current political climate .[sentence: 3,-4] [result: max + and - of any sentence][overall result = -1 as pos<-neg] (English)

Fig. 1. Example Sentence analysis

3.2 Implementation

3.2.1 *implementation stack.*

- python 3.7 as programming language
- pandas for the data handling
- Sentistrength[Thelwall et al. 2010] for the sentiment analysis
- classic python libraries

3.2.2 implementation setup. The given file[Kolhatkar et al. 2018] is separated into two files. One file contains all statements the second file contains all comments. Both files contain the column article_id. Based on this article_id the statement and comments can be mapped together afterwards. Our algorithm evaluates with Sentistrength the statement and the comments separately and stores the result. In the following we refer to the analysed files as data, this data will be stored in new csv files. So that the data can be reused without doing the sentiment analysis all over again. Afterwards we read the data file again and do some data preparations. Then we analysed the data, the results are provided in section 4

3.2.3 problems. During the implementation we faced a few problems. First of all the date handling was quite complicated, because we needed some data formatting and preparation. Moreover we were not able to use the Sentistrength[Thelwall et al. 2010] method in parallel. Because of this issue the analysis did take 4 days for 673.512 texts. Next time we would choose a smaller dataset for experiments of this kind.

4 GROUP DECISION BASED ON COMMENT REACTIONS

Group decisions are never easy, especially if the discussion does not take place in person. There exist many different ideas and methods to come to a decision[Morente-Molinera et al. 2019]. Most of the methods we found depend on a revisiting process. Meaning that a statement is given and the discussion starts, based on this discussion alternatives and renewed statements are given. This process can repeated as long as a certain agreement level is achieved [Dong et al. 2018]. On social media this process can be difficult, because the statements are posted separately. The Authors[Morente-Molinera et al. 2019] suggests to map different topics by there Hashtags.

In our case, we will evaluate how much agreement and disagreement an article gets by analysing the comments below a statement. We are summing up all positive and negative reactions of each comment with Sentistrength[Thelwall et al. 2010]. Furthermore, we want to experiment if a positive statement gets more positive comments than a negative statement. This means that we analyse the statement itself as well. Each sentence gets analysed by itself and then the results will be summarised together per statement. We are aware that this method is probably not sufficient enough to make a general assumption of the articles opinion.

4.1 Test results

In the following chapter we will present some of our results based on the Sentistrength[Sentistrength 2020] analysis of the SFU-Opinion and Comments Corpus[Kolhatkar et al. 2018] information. For better accuracy we did not consider statements that had less than 10 comments, this leads to $6,427$ statements with an average of 102 comments. The number of comments are balanced over the years, there was no increase or decrease visible. The sentiment range is always between -5 and 5 even though in most of the figures only -3 to 3 is shown, because there are no high negative or positive values.

Figure 2 shows the distribution of the sentiment values for all statements. We are able to see that the overall statements are sightly negative by a medium sentiment value of -1. Interestingly the Sentistrength[Thelwall et al. 2010] does not rate a statement higher than 3 or lower than -3. There are only some outliers, most other statements are located between -2 and 1.

In Figure 3 we see the rate of agreement of all comments with their corresponding statement. A value of 50% is defined as sentiment value of 0 (= neutral). The median of 45.66% leads to the conclusion that the comments are relatively balanced. This can be explained by two scenarios: (A) Many strongly positive and many strongly negative comments. (B) Most comments by themselves are neutral. To find out the more likely scenario we calculated the standard deviation with a value of $1,2$. Therefore, we conclude that each comment is neutral by itself.

We used the Pearson correlation coefficient to figure out if there is a correlation between the sentiment of the statements and the sentiments of their comments. According to our calculation the Pearson correlation coefficient is 0.263 which is a low positive correlation. This means more positive statements get slightly more positive comments. A graphical visualisation of this is visible in Figure 4. For completeness we did the Spearman correlation coefficient as well and received similar results, with a value of 0.267.

Figure 5 shows the discussion time per statement from 2013 to 2016. The median discussion time was 3 days. For this calculation we considered the date of the first comment and the date of the comment at the 90% mark of all comments of the corresponding statement. We did choose to only consider 90% of the comments to avoid random comments years later. As you can see in the Figure 5 some discussions took way longer than the average. Interestingly, since 2016 the reaction time decreases, as visible in Figure 5. Since the data[Kolhatkar et al. 2020] have been collected for at least 2 years per article, also comments after the last publishing date has been considered. Therefore, the data collection should not be a significant reason for the decline in discussion time.

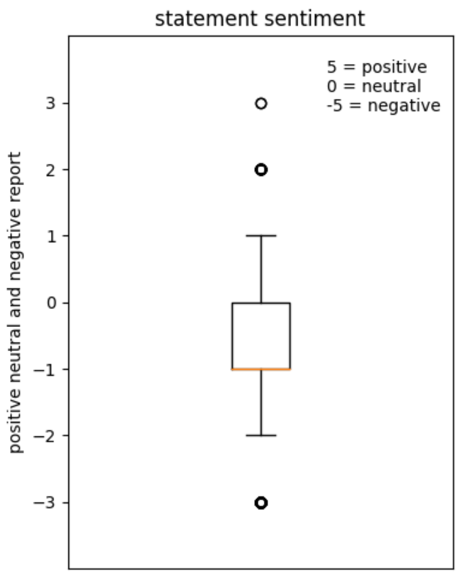

Fig. 2. Overall statement sentiment

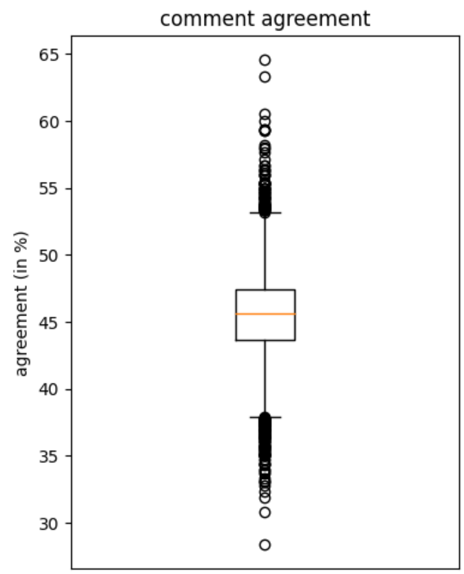

Fig. 3. Overall agreement on the articles

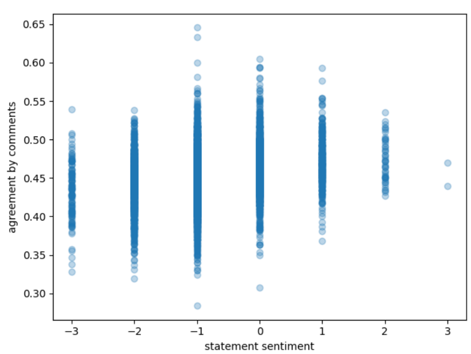

Fig. 4. correlation statement sentiment and comment agreement

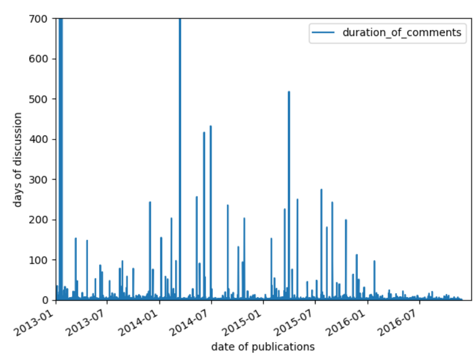

Fig. 5. discussion time

5 CONCLUSION AND OPEN ISSUES

The results were not what we expected, especially because the [Kolhatkar et al. 2020] did already choose opinion articles for their corpus, we assumed that there would be more variety in the sentiments of the statements. Moreover we assumed that the comments would tend to a more negative or positive. Nevertheless in our experiment it becomes clear that most of the statements and comments are neutral, but tend to be slightly negative. We think that it is possible to conclude that a statement's sentiment has an minor positive effect on the sentiment of its received comments. Therefore if a statement's sentiment is positive it receives rather positive comments as well and vice versa.

Further research can be done in the following areas: The dataset[Kolhatkar et al. 2018] provides more information for further experiments. It would be possible to categorise the statements to certain topics. Furthermore, the comment itself gets reactions by votes and replies, which we would find interesting to consider. In our evaluation we ignored statements that had less than 10 comments, it would be interesting to know why the left out 3912 statements did receive less comments. Furthermore, the accuracy of the sentiment results would be interesting.

REFERENCES

Alessia D'Andrea, Fernando Ferri, Patrizia Grifoni, and Tiziana Guzzo. 2015. Approaches, Tools and Applications for Sentiment Analysis Implementation. *International Journal of Computer Applications* 125 (09 2015), 26–33. https://doi.org/10.5120/ijca2015905866

Yucheng Dong, Quanbo Zha, Hengjie Zhang, Gang Kou, Hamido Fujita, Francisco Chiclana, and Enrique Herrera-Viedma. 2018. Consensus reaching in social network group decision making: Research paradigms and challenges. *Knowledge-Based Systems* 162 (2018), 3 – 13. https://doi.org/10.1016/j.knosys.2018.06.036 Special Issue on intelligent decision-making and consensus under uncertainty in inconsistent and dynamic environments.

Doaa Mohey El-Din Mohamed Hussein. 2018. A survey on sentiment analysis challenges. *Journal of King Saud University - Engineering Sciences* 30, 4 (2018), 330 – 338. https://doi.org/10.1016/j.jksues.2016.04.002

Varada Kolhatkar, Hanhan Wu, Luca Cavasso, Emilie Francis, Kavan Shukla, and Maite Taboada. 2018. SFU Opinion and Comments Corpus. (2018). https://doi.org/10.25314/71358419-06df-4b07-a160-7913737ca28f

Varada Kolhatkar, Hanhan Wu, Luca Cavasso, Emilie Francis, Kavan Shukla, and Maite Taboada. 2020. The SFU Opinion and Comments Corpus: A Corpus for the Analysis of Online News Comments. *Corpus Pragmatics* 4, 2 (01 Jun 2020), 155–190. https://doi.org/10.1007/s41701-019-00065-w

Isa Maks and Piek Vossen. 2012. A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems* 53, 4 (2012), 680 – 688. https://doi.org/10.1016/j.dss.2012.05.025 1) Computational Approaches to Subjectivity and Sentiment Analysis 2) Service Science in Information Systems Research : Special Issue on PACIS 2010.

J.A. Morente-Molinera, G. Kou, K. Samuylov, R. Ureña, and E. Herrera-Viedma. 2019. Carrying out consensual Group Decision Making processes under social networks using sentiment analysis over comparative expressions. *Knowledge-Based Systems* 165 (2019), 335 – 345. https://doi.org/10.1016/j.knosys.2018.12.006

C. Roy, R. Dey, C. Chaudhuri, and D. Das. 2019. Emotion Predictor Using Social Media Text and Graphology. In *2019 IEEE 9th International Conference on Advanced Computing (IACC)*. 96–102. https://doi.org/10.1109/IACC48062.2019.8971577

Hassan Saif, Yulan He, Miriam Fernandez, and Harith Alani. 2016. Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management* 52, 1 (2016), 5 – 19. https://doi.org/10.1016/j.ipm.2015.01.005 Emotion and Sentiment in Social and Expressive Media.

K. Schouten and F. Frasincar. 2016. Survey on Aspect-Level Sentiment Analysis. *IEEE Transactions on Knowledge and Data Engineering* 28, 3 (2016), 813–830. https://doi.org/10.1109/TKDE.2015.2485209

Sentistrength. 2020. *SentiStrength*. Retrieved December 10, 2020 from http://sentistrength.wlv.ac.uk/

Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology* 61 (12 2010), 2544–2558. https://doi.org/10.1002/asi.21416