Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

Sentimental Analysis of Olympics Tweets

Dr. Rajeev Srivastava, Faculty UPES Dehradun India
Dr. Atul Kumar Agarwal, Professor, Sushant University – Gurgaon
Dr. GunjanA.Rana, Associate Professor , Sushant University – Gurgaon

ABSTRACT
The impact of social media on business, entertainment, education, and in almost every sector is increasing day by day. Peoples are expressing their positive as well as negative emotions on different platforms of social media. On Twitter also peoples are sharing their emotions through tweets. Peoples are retweeting on both positive as well as negative types of tweets. The objective of this paper is to study the sentiments of upcoming Olympics tweets to understand the emotions of people and their impact on retweets. The data for the study was collectedfrom the Twitter website in the form of tweets related to the upcoming Olympics.Sentimental analysis of tweets was used using python library textblobto understand the feelings or emotions of people. The results of the study indicated that the tweets having lower polarization are getting more tweets.This study helps the organizations to understand that the negative emotions (tweets) are more supported in social media especially on Twitter.

Key Words – Sentimental analysis, Tweets, Olympics, Retweet, Polarity, Subjectivity

1. Introduction

Social media text data deliver people's opinions, thoughts, and emotions. Mining of this textual data or unstructured data to understand the opinion, thoughts, and emotions of a group of peoples is also known as opinion mining or sentimental analysis.From the last decade, social media data has been the main source for spreading information in various domains like entertainment, politics, andbusiness.The main source of these social media platforms are Twitter, Facebook, and Instagram. These platforms are using unstructured as well as structured data for that purpose.This growing volume of data introduced new avenues to know the insight of the different issues or topics specifically in text data.

No study is available in the literature to identify what is the impact of a retweet on positive or negative tweets especially in the field of sports. So, it’simportant to know whether positive tweets are getting more retweets or the negative tweets are getting more retweets. So the main objective of this paper is to understand the sentiment of people's emotions of upcoming Olympics tweets. To identify the polarity and subjectivity of tweets and finally to understand the impact of retweets on polarization and subjectivity.

API is used for getting the tweets from Twitter. Use Natural Language Processing (NLP) packages of python to extract things like polarity and subjectivity of the tweets. Python library

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

textblob is used for the processing of textual data. The result shows that as the retweets count (retwc) is increasing the polarity is going down. Which shows higher retwc have lower polarization. This means negative comments are having more retweets.
In the second part review of literature in terms of sentimental analysis is presented. In the third section, the purpose or the objective of the study is discussed. In the fourth part, the methodology used to find the answer tothe research question is presented. In the rest of the part discussion, the limitation and conclusionsof the study are discussed.

2. Review of Literature

In the last decade, social media has evolved as an important driver for spreading and acquiring different types of information in different sectors. The major sectors are business (Beier & Wagner, 2016), entertainment (Shen, Hock Chuan, & Cheng, 2016), science (Chen & Zhang, 2016), crisis management (Hiltz, Diaz, & Mark, 2011; Stieglitz, Bunker, Mirbabaie, &Ehnis, 2017a) and, politics (Stieglitz & Dang-Xuan, 2013).Social media is generating a variety of data, which can be categorized as unstructured data and structured data (Baars & Kemper, 2008).The textual content is an example of unstructured data, while the friend/follower relationship is an example of structured data. Based on the mental procedures the confidence is characterized into influence and assessment. As far as the research on confidence there is no cultural research evidence is that is completed (Luqman N. 2020) .
The continuous growth in the usage of social media opens new avenues for identifying the pattern in the communication like this data can be analyzed to know the insight into the trends, issues, and even the influential actor of the communication. Many studies havealready conducted to know how people mood’s changes from morning to evening, on weekdays and season-wise on specific subjects or issues (Golder and Mac, 2011, Kim, Choi, & Natali, 2016; Li & Huang, 2014; Oh, Hu, & Yang, 2016).The Android Permission Model should keep the normal permission in a signature or dangerous category as there can be malicious activity which can be misused(Malik, S., &Khatter, K. 2016).

No study is available in the literature to identify what is the impact of a retweet on positive or negative tweets especially in the field of sports. So, it’s important to know whether positive tweets are getting more retweets or the negative tweets are getting more retweets. So the main objective of this paper is to understand the sentiment of people's emotions of upcoming Olympics tweets. To identify the polarity and subjectivity of tweets and finally to understand the impact of retweets on polarization and subjectivity. In this background, it is hypothesized that:

H1: Retweets havean association with the polarity of tweets.

H2: Retweets havean association with the subjectivity of tweets.

3. Objectives of the Study

- To identify the polarity and subjectivity of tweets.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

- To understand the impact of a retweet on polarization and subjectivity
- To open new avenues for future research.

4. Research Methodology

The study was Casual and the API was used for data collection. Thepopulation for the research study was all the tweets related to the upcoming Olympics 2021. Individual tweets of the upcoming Olympics are treated as sample elements and the sample size of the study was 45tweets. The tweets dated 9 August to 17 August, 2020was used to identify respondentsfor the study.

Natural Language Processing (NLP) packages of python are used to identify the polarity and subjectivity of each tweet.Polarity is used to extract positive, negative, or neutral emotions of each tweete.g.(+,-,0) from the textual data.

Subjectivity is the related concept and used to extract whether people's opinion is objective or subjective about that textual data. Normally its value lies between 0 to 1. Where values less than 0.5 indicate objectivity and a value greater than and equal to 0.5 indicate subjectivity.

The details of username having retweets count greater than 100 are shown in Fig. 1. This indicates that the username ‘Olympics’ has the highest number of retweets (1,837) followed by ‘David Rudisha’ (1,751) and so on. The details of followers for each username is shown in Fig. 2. This indicates that the username ‘CNN’ has the highest number of followers (99,067,846) followed by ‘Olympics’ (36,196,428) and so on.

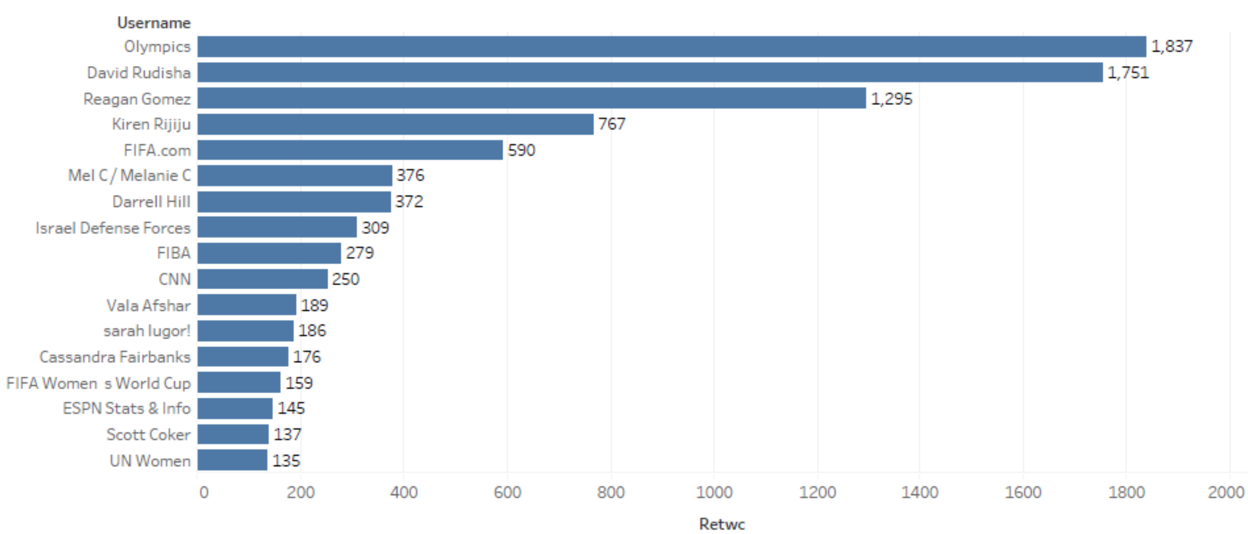

Fig1. Retweets for each Username

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

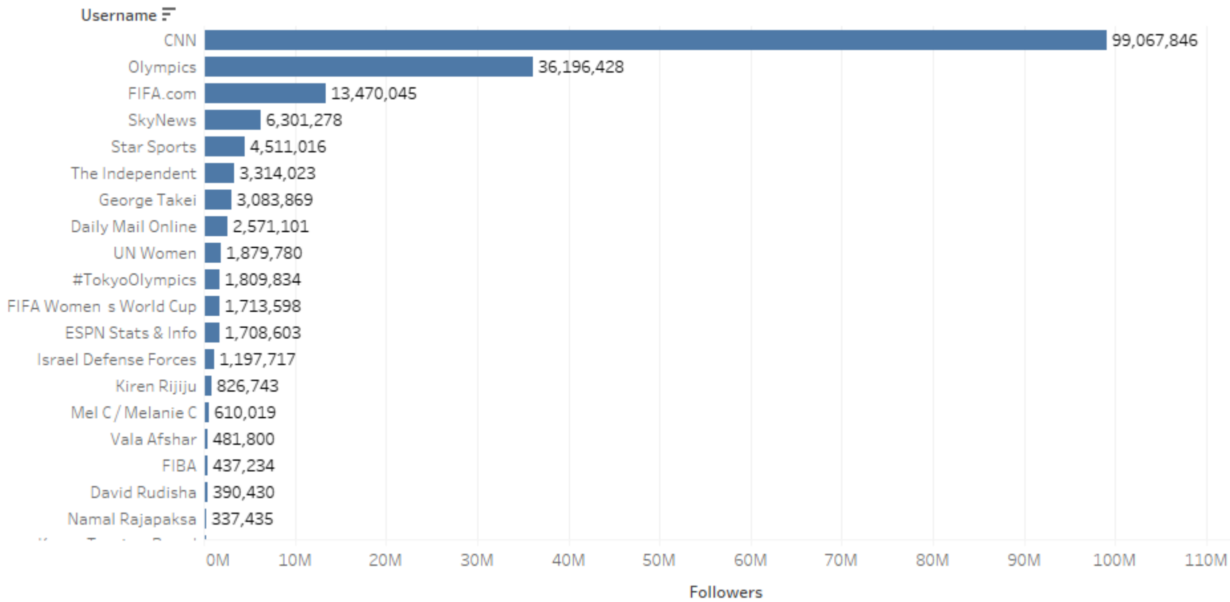

Fig2.Followers for each username

The average of positive and negative Polarity is shown in Fig 3. This indicates that the average of positive polarity is .38, while the average of a negative polarity is .21. Which means that the number of positive tweets is more as compared to negative tweets.The average of objectivity and Subjectivity is shown in figure 4. The average value of subjectivity is .68 and the average value of objectivity is .22. These values indicate that the count of subjective tweets is more as compared to objective tweets.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

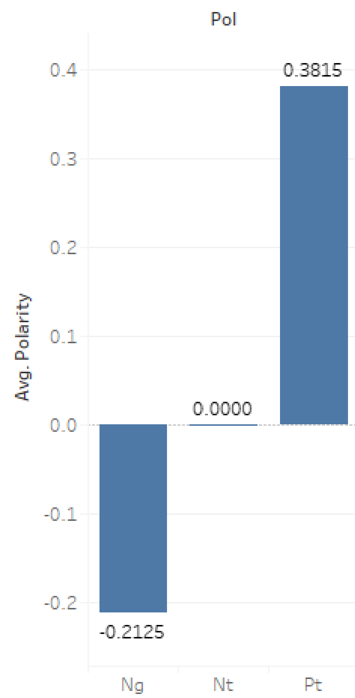

Fig3. Average of positive and negative Polarity

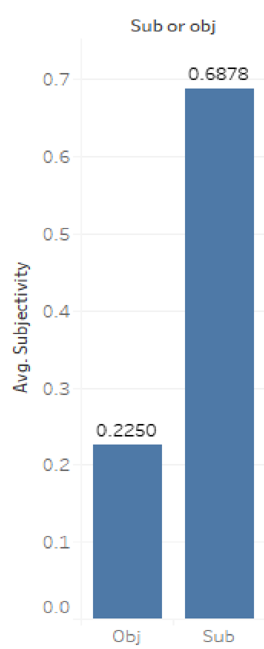

Fig4. Average of objectivity and Subjectivity

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

Python library textblobis used for the processing of textual data. From the textblob package,TextBlob class is specifically used for this purpose.Polarity values lie between -1 to 1 and subjectivity value lies between 0 to 1.From each tweet derive to access polarity as well as subjectivity. Normally polarity value greater than 1 is considered positive, a value less than 1 considered negative, and a value equal to zero is considered neutral. In the case of subjectivity value greater than 0.5 considered tweet is subjective and value less than or equal to 0.5 considered objectives.

	author_ id	retwc	followers	friends	polarity	subjectivity
author_ id	1.000000	0.142682	-0.151641	-0.069355	0.074764	0.319723
retwc	0.142682	1.000000	-0.032962	-0.111402	-0.200128	0.037137
followers	-0.151641	-0.032962	1.000000	-0.087200	-0.156548	-0.300552
friends	-0.069355	-0.111402	-0.087200	1.000000	0.235431	0.133054
polarity	0.074764	-0.200128	-0.156548	0.235431	1.000000	0.500765
subjectivity	0.319723	0.037137	-0.300552	0.133054	0.500765	1.000000

Fig. 5. Correlation betweenretwc and polarity

The correlation matrixis shown in figure 5. In this matrix, the correlation between retweet count (retwc) and polarity value is -0.2. The result shows that as the retweets (retwc) are increasing the polarity is going down. The scatter plow of retweet count (retwc) and polarity shown in figure 5 shows higher retwc have lower polarization. This means negative tweets are having more retweets.

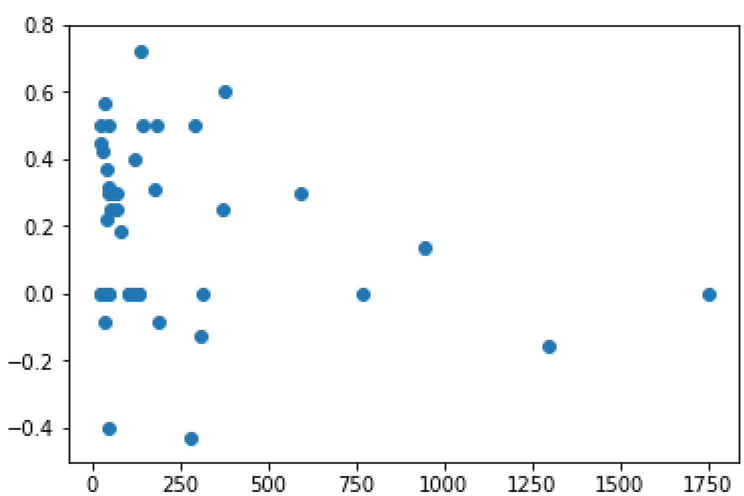

Fig 6. Scatterplot of retweet count and polarity

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

In figure 5 the value of the correlation between retwc and subjectivity is.037, Which shows that there is no correlation between retwc and subjectivity. While extracting tweets having subjectivity value >0.5 or which is subjective and then again find the correlation between retwc and subjectivity, which shown in figure 7.. Now the correlation between retwc and subjectivity has a positive correlation(.64) which was earlier (.037). This .64 correlation value indicates that the more subjective the tweet is more will be the no of retweets.

	author_ id	retwc	followers	friends	polarity	subjectivity
author_ id	1.000000	0.661887	-0.215020	-0.099641	-0.271619	0.384706
retwc	0.661887	1.000000	0.069957	-0.248787	-0.356118	0.640650
followers	-0.215020	0.069957	1.000000	-0.150302	0.189727	0.026811
friends	-0.099641	-0.248787	-0.150302	1.000000	0.222615	0.064946
polarity	-0.271619	-0.356118	0.189727	0.222615	1.000000	0.327015
subjectivity	0.384706	0.640650	0.026811	0.064946	0.327015	1.000000

Fig. 7. Correlation betweenretwc and subjectivity

5. Discussion of Results

The main objective of this study was to identify the polarity and subjectivity of tweets related to the upcoming Olympics and to understand the impact of a retweet on polarization and subjectivity.The study reveals that the average positive polarity value 0.38 and the average negative polarity value 0.21 shown in figure 3. Indicates that the number of positive tweets is more as compared to negative tweets. Similarly the average value of subjectivity .68 and the average value of objectivity .22 shown in figure 4. indicates that the count of subjective tweets is more as compared to objective tweets.

The study found that there isa negative cause and effect relationship between retweet count and polarity. This means negative tweets are having more retweet counts.The study also found that there isa strong and positive cause and effect relationship between retweet counts and subjective tweets. This .64 correlation value indicates that the more subjective the tweet is more will be the no of retweets.

6. Limitations of the Study

The sample size of the current study was 45 tweets. So it's not justified to generalize that tweets with negative polarity are having more tweets. Generalizations would be further justified by

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 24, Issue 2, 2020, Pages. 427 - 435
Received 24 October 2020; Accepted 15 December 2020

using a bigger sample size.This research was conducted using tweets related to the upcoming Olympics therefore, the resultsof this study cannot be generalized for tweets related to other sectors. The tweets collected for a certain period related to the sentiments of the people. Sentiments keep on changing, so accordingly result will also modify.

7. Conclusion

The study shows that the number of positive tweets is more as compared to negative tweets related to the upcoming Olympics. It has also been found that subjective tweets are more as compare to objective tweets. During the analysis, it has been found that negative tweets are having more retweets and subjective tweets get more retweets as compared to the objective ones.

References

8. Beier, M., & Wagner, K. (2016). Social media adoption: barriers to the strategic use ofsocial media in SMEs. Proceedings of the European conference on information systems.

9. Shen, Y., Hock Chuan, C., & Cheng, S. H. (2016). The Medium Matters: Effects on WhatConsumers Talk about Regarding Movie Trailers. In Proceedings of the International
10. Conference on Information Systems.
11. Chen, F., & Zhang, L. (2016). How to integrate social media in IS curriculum, especiallyfor a small IS program? Proceedings of the Americas conference on information systems.

12. Hiltz, S.R., Diaz, P., & Mark, G. (2011). Social media and collaborative systems for crisismanagement, ACM Transactions on Computer-Human Interaction, Vol. 18 Issue 4,December 2011, ACM New York, NY, USA, DOI: https://doi.org/10.1145/2063231.

13. Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: A socialmedia analytics framework. Social Network Analysis and Mining, 3(4), 1277–1291.http://dx.doi.org/10.1007/s13278-012-0079-3.

14. Stieglitz, S., Bunker, D., Mirbabaie, M., & Ehnis, C. (2017a). Sense-Making in Social MediaDuring Extreme Events. Journal of Contingencies and Crisis Management (JCCM).http://dx.doi.org/10.1111/1468-5973.12193.

15. Baars, H., & Kemper, H.-G. (2008). Management support with structured and unstructureddata – An integrated business intelligence framework. Information SystemsManagement, 25(2), 132–148. http://dx.doi.org/10.1080/10580530801941058.

16. Golder, S. A., & Macy, M. W. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep,and Daylength Across Diverse Cultures. Science, 333(6051), 1878—1881. https://doi.org/10.1126/science.1202775.

17. Kim, Y., Choi, K. S., & Natali, F. (2016). Extending the Network: The Influence of OfflineFriendship on Twitter Network Full papers. Proceedings of the Americas Conference onInformation Systems.

18. Li, Z., & Huang, K.-W. (2014). The Monetary Value of Twitter Followers: Evidences fromNBA Players. In Proceedings of the International Conference on Information Systems.

19. Luqman N. (2020) Relationship between Religiosity, Self-Esteem and Locus of Control: A Comparative Study. J Psychiatry Psychol Res, 4(1): 204-209.

20. Malik, S., &Khatter, K. (2016). BEHAVIOUR ANALYSIS OF ANDROID APPLICATION. International Journal of Control Theory & Application, 9(10), 358-372.

21. Oh, C., Hu, H.-f., & Yang, W. (2016). Social Media Information Diffusion and EconomicOutcomes: Twitter Retweets and Box Office. In PACIS 2016 Proceedings.

22. Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: A socialmedia analytics framework. Social Network Analysis and Mining, 3(4), 1277–1291.http://dx.doi.org/10.1007/s13278-012-0079-3.

23. Bruns, A. (2013). Faster than the speed of print: Reconciling ‘data’ social media analysisand academic scholarship. First Monday, 18(10), s1. http://dx.doi.org/10.5210/fm. v18i10.4879.