Lexicon-based Sentiment Analysis and Emotion Classification of Climate Change Related Tweets

Temitayo Matthew Fagbola[1,2][0000-0001-6631-1002], Abdultaofeek Abayomi[3][0000-0003-3129-5246], Murimo Bethel Mutanga[4][0000-0002-6021-8613], and Vikash Jugoo[5]

[1]Department of Computer Science, Federal University Oye, Oye-Ekiti, Ekiti State, Nigeria
[2]Institute for System Sciences, Durban University of Technology, Durban 4000, South Africa
[3,4,5] Department of Information and Communication Technology, Mangosuthu University of Technology, P.O. Box 12363, Jacobs 4026, Durban, South Africa

temitayo.fagbola@fuoye.edu.ng

Abstract. The concerns for a potential future climate jeopardy has steered actions by youths globally to call the governments to immediately address challenges relating to climate change. In this paper, using natural language processing techniques in data science domain, we analyzed twitter micro-blogging streams to detect emotions and sentiments that surround the Global youth Climate Protest (GloClimePro) with respect to #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime hashtags. The analysis follows tweet scrapping, cleaning and preprocessing, extraction of GloClimePro-related items, sentiment analysis, emotion classification, and visualization,. The results obtained reveal that most people expressed joy, anticipation and trust emotions in the #ThisIsZeroHour and #ClimateJustice action than the few who expressed disgust, sadness and surprise. #ClimateJustice conveys the most positive sentiments, followed by #ThisIsZeroHour and the #WeDontHaveTime . In all the evaluations, a considerable number of people express fear in the climate action and consequences. Thus, climate change stakeholders and policy makers should consider the sentiments and emotions expressed by people and incorporate such outcomes in their various programmes toward addressing the climate change challenges especially as it affects the ecosystem.

Keywords: classification, climate change, ecosystem, emotion, nature, sentiment

1 Introduction

Climate change remains a global concern. Unlike the past two decades when a need was desired to protect the environment, as it stands today, many people across the globe have started noticing the consequences of climate change. Sustained climate change has been projected by the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report to have dire and irreversible effects for the people as well as the ecosystem across the globe. The negative impacts of the climate change include declining crop yields, drought, flood, wildfires, ocean acidification, increased frequency and intensity of extreme daily temperatures [1]. In order to make nations more serious about climate change and move from business-as-usual situations to transformative change in various mitigation paths, a lot of awareness campaigns including at the recent 2021, 26th Conference of the Parties to the United Nations Framework Convention on Climate Change (COP26), held at Glasgow, Scotland; and protests have taken place across the globe, urging the nations to act now towards a safe nature, safe earth and a safe future. Most of these campaigns and protests are often organised online and undertaken by the youths as they constitute a significant 2.2 billion of the world's population. As a result, many countries are in the process of developing adaptation plans and disaster risk reduction strategies to help people adapt to changing climatic conditions wherever they live. Then, it becomes imperative for the climate change stakeholders to understand the distribution of sentiments and emotions that the public share regarding climate change. This is in a bid to better understand public opinions so that a collective success at fighting climate change could be achieved. With traditional sentiment analysis, a comprehensive interpretation of the contextual meanings associated with a text is limited. Yet, emotion mining approaches can be applied to gain more in-depth understanding of the user's perspective about an event. Such approaches extract emotion of users from some given texts [2]. Sentiment and emotion classification methods can be categorized into the lexicon, the machine learning and the rule-based approaches [2]. However, most recent works have strongly emphasized the effectiveness of lexicon-based approach due to its simplicity. Among other important insights, a major research question is: “What is the distribution of public sentiments and emotions regarding climate change actions globally?”. In this paper, we leveraged Twitter micro-blogging streams and word-emotion association lexicon tools to detect emotions and sentiments that surround the Global youth Climate Protest (GloClimePro) with respect to #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime hashtags. The rest of the paper is organized as follows: Section 2 reviewed related work regarding climate change and sentiment analysis. Section 3 presents the materials and method while Section 4 presents the results and discussion and in Section 5, the conclusion is presented.

2 Literature Review

Some studies that have been conducted on climate change using the sentiment analysis and topic modeling techniques are hereby presented. The authors [3] determined public awareness of the terms climate change and global warming by assessing the relative search volume (RSV) patterns referencing the terms “climate change” and “global warming” from a total of 21,182 and 26,462 tweets, respectively. Sentiment analysis was performed on the dataset using the Semantria® software while the differences in RSV’s for the terms are identified using a paired t-test. In another study [4], public opinion on climate change was evaluated from 390,016 tweets associated with keywords like carbon dioxide, fossil fuel, carbon footprint and emissions. The tweets were analyzed by applying the Valence Aware Dictionary and sEntiment Reasoner and the Latent Dirichlet Allocation topic modeling technique. The study found that the overall discussions on the issue of climate change is negative particularly with respect to extreme weather conditions, government and institutional policy issues.

The authors in [5] conducted a research tagged “MC3 Meeting the Climate Change Challenge” over a six-year period to examine climate change innovations and non-computational public sentiment in eleven Canadian local municipalities. Due to the diverse sentiments surrounding the word “change”, for instance, as it relates to its urgency, cost, consequences and lifespan, various drivers and barriers to change were exposed. An openended, semi-structured interview that drew upon multi-level perspectives, social practice theory and social-ecological systems from across British Columbia, Canada was used. The positive changes and sentiments recorded are connected with the quality of the local government leadership and staff which exist in the institutions. However, the negative sentiments observed are connected with the magnitude and swiftness of expected change.

Abbar et al. 2017 [6] analysed 109.6 million climate change-related tweets and daily weather conditions data to understand public gravitation in relation to topics including politics, economy, energy, air quality, disasters and sandstorm. By employing content and correlation analyses, some of the findings obtained indicate that people’s interests are driven by widely covered events that have direct link to their daily lives and that the number of users engaging in climate change discussion is not necessarily in an increasing trend. In the study conducted by [7], over 4.5 million English language tweets, relating to the 2015’s 21st Conference of the Parties (COP21) of the United Nations on Climate Change held in Paris, was analysed based on spatiotemporal distribution of emotions, opinions and demographics in order to discover how the various users discuss, behave and form opinions regarding the climate change event. The tweets discourse were categorised into nine groups which include energy, weather, economy, agriculture/forestry, water, security, climate denial, air issues and animals. As parts of their findings, it was discovered that organisations express the most anger and joy emotion laden tweets toward water, the most disgust tweets is toward food, the most fear tweets toward security, the most sadness tweets toward animal, the most surprise tweets toward climate denial, while the most negative opinions was toward air issues, the most positive toward water and the most neutral toward food.

Twitter data was used to complement insights into climate change perceptions in a study by [8]. Time series and sentiment analysis techniques were applied to 494,097 tweets relating to climate change in order to detect sentiments associated with climate-related hazards such as wildfire, hurricane, flood and drought. Naïve Bayes (NB) and support vector machines algorithms were applied for the tweet sentiment classification task. The NB achieved the best performance of 76.54% accuracy, F1 score of 0.8057; and 76.77% accuracy, F1 score

0.8278 for both the subjective and polarity classification tasks, respectively. Ojala and Lakew [9] focused on the youth as a crucial group to mobilise and integrate into the efforts of curbing climate change since they would be societies' leaders in the future. The study investigated how young people are influenced by communications and how they use the social media to interact about climate change. The authors recommended the art-based approaches as vital ways to communicate with the young people about climate change.

3 Materials and Method

The lexicon-based analytic framework used for the detection of sentiments and emotions that surround the Youth Climate Change Action is presented in Figure 1. The framework, which was originally developed by [2], is a five-stage architecture that is made up of Twitter data acquisition, data cleaning, data preprocessing, the emotion and sentiment analytic engine, and visualization.

3.1 Twitter Dataset Acquisition

Tweets between 01 September 2019 and 28 June 2020, in connection with the global youth climate action on Twitter, were collected using the Twitter REST API into a separate Microsoft Excel file for each distinct hashtag. The three most prominent hashtags used include #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime. As presented in Table 1, 10120 tweets and 41,077 retweets were obtained using #ThisIsZeroHour. Similarly, 19,449 tweets and 144,484 retweets are connected with #ClimateJustice. The #WeDontHaveTime produced 6706 tweets and 36172 retweets. Samples of tweets from #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime are shown in Tables 2, 3 and 4.

3.2 Data Cleaning and Pre-processing

A lot of noise is often associated with Twitter data. Hence, data cleaning becomes highly imperative. Twitter handles, stopwords, hyperlinks, special and repeated characters in tweets, duplicate hashtags and tweets, empty spaces and retweet entities were all removed.

Furthermore, bots detection and removal, tokenization, stemming and part-of-speech tagging were conducted. Bots contained in the Twitter data were detected using the Botometer [10]. With a bot score of 0.6, 1063 tweets from the #ThisIsZeroHour, 2,207 tweets from the #ClimateJustice and 886 tweets from the #WeDontHaveTime originated from bot accounts and were outrightly removed. We applied the *gettokens* () in the Syuzhet library to tokenize the remaining tweets into a set of unigrams. The *stemDocuments* () in the "*tm*" library within the *R* package was applied to reduce each unigram to its shortest form. Using the part-of-speech tagging, we disambiguate the meaning and the lexical categories associated with each unigram.

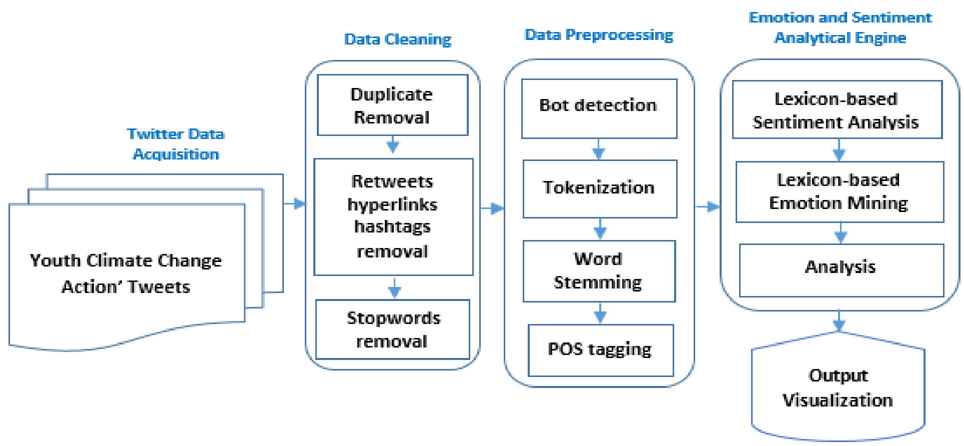

Fig. 1. The lexicon-based sentiment and emotion analytic framework.

Table 1. Youth climate action hashtags and associated tweets.

Hashtags	Number of Tweets	Number of retweets
#ThisIsZeroHour	10,120	41,077
#ClimateJustice	19,449	144,484
#WeDontHaveTime	6,706	36,172

3.3 Lexicon-based Emotion and Sentiment Analytic Engine

In this paper, the word emotion association lexicon (emolex) dictionary and the Syuzhet library functions are used to detect and estimate polarity of sentiments (positive and negative) and emotions (anger, anticipation, disgust, fear, joy, sadness, surprise and trust) expressed in the global youth climate action tweets. The lexicon-based emotion classification process is presented in Algorithm 1. We define emotions from tweets as $E = e_1, e_2, \dots, e_8$ based on the Plutchik's wheel of emotions [11]. The set of tweets whose emotions are to be identified is defined as $S = s_1, s_2, \dots, s_n$ while the number of times that a word appears in a tweet is denoted in Equation 1 as:

$$E(e_i|w_x) = \sum_{s \in S} E(e_i|S) * F_s(W_x) \tag{1}$$

where $\mathrm{E}(\mathrm{e_i}|\mathrm{S})$ the emotion contained in the tweet, $\mathrm{S}.\,\mathrm{F_s}(\mathrm{e_i})$ is a pointer function set to 1 if $\mathrm{e_i} \in \mathrm{s}$ and 0 otherwise

Table 2. Sample tweets from #ThisIsZeroHour hashtag.

10020	We live in a sad society where imbeciles love & care for Dogs & Cats more than their children
10024	PS/ @ThisIsZeroHour pic.twitter.com/1y9QXwFHsW
10025	we'll be sending out more info in terms of location and time in a bit!
10026	Where are all those #ClimateStrikers on Twitter? Everybody should tweet about #Climate Crisis!

Table 3. Sample tweets from #ClimateJustice hashtag.

315	"My activism for racial justice and activism for climate justice are one in the same"
317	Don't miss "The Youth Behind the Movement: A climate Activism Webinar"! 6 youth climate activists in the Cowichan Valley who found community in climate action
318	Jumping on the bandwagon What has Sinn Fein ever done to address climate Justice? Answer on a postcard!

Table 4. Sample tweets from #WeDontHaveTime hashtag.

2617	Amazing…!You still haven't figured out that there literally is NO "climate crisis" or emergency! What's it like abandoning all things scientific?
2618	Many news stories in media because of this. Please help us share this method of creating awareness!
2619	The link between nature and humanity cannot be broken.# WeDontHaveTime #ForNature #WorldEnvironmentDay
2620	pic.twitter.com/MKsfqylzZa

Algorithm 1: Step by step process for **l**exicon-based emotion classification of tweets

$\boldsymbol{Input}: S - a\ sample\ set\ of\ tweets;\ L_w - a\ lexicon\ word\ dictionary\ of\ emotions;$
$\quad F - a\ pointer\ function;$
$\boldsymbol{Required}: W_x - a\ number\ of\ times\ the\ word\ L_w\ occurs\ in\ a\ given\ text;$
$\quad E - emotions\ in\ Tweets$
$for\ all\ i\ to\ n\ do$
$\quad if\ L_w[i] == S[i] then\ F \leftarrow 1\ else\ I \leftarrow 0 \quad end\ if$
$end\ for$
$for\ all\ i\ to\ N\ do\ \ F * W_x\ \ end\ for$

$for\ all\ i\ to\ N\ do$

$\quad \sum_{s \in S} E * F$

$end\ for$

4 Results and Discussion

The results of sentiment and emotion classification obtained for #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime, the three most significant hashtags for the global youth climate action, are hereby presented. Wordcloud visualizations, showing the metadata of the hashtags and close keywords from connected tweets, for the #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime hashtags are presented in Figures 2, 3, and 4, respectively. Important insight drawn from the wordclouds for #ThisIsZeroHour are words such as “Climate”, “Change”, “Sunrisemovement”, “People”, “Youth” and “Jamie Margolin”. Jamie Margolin is an American climate justice activist and the co-executive director of Zero Hour [12]. “Climate” and “Justice” are the two major words boldly represented in the wordcloud for #ClimateJustice. Also, “Climate” and “Climate change” are recognized in the wordcloud for #WeDontHaveTime. The emotion scores for the #ThisIsZeroHour are 4001, 5509, 1397, 4908, 4951, 2489, 3329, 6550, 6229 and 13020 for anger, anticipation, disgust, fear, joy, sadness, surprise, trust, negative and positive respectively as presented in Figure 5. #ThisIsZeroHour conveys 67.6% positive and 32.4% negative sentiments indicating that most people are happy with the movement. Emotions associated with the #ThisIsZeroHour include disgust, sadness, surprise, anger, fear, joy, anticipation and trust in increasing order of polarity score. By implication, most people expressed joy, anticipation and trust in the #ThisIsZeroHour action than the few who expressed disgust, sadness and surprise. The emotion scores for the #ClimateJustice hashtag are 10002, 14018, 3000, 11900, 12200, 7000, 6000, 39000, 17000 and 56000 for anger, anticipation, disgust, fear, joy, sadness, surprise, trust, negative and positive respectively as shown in Figure 6.

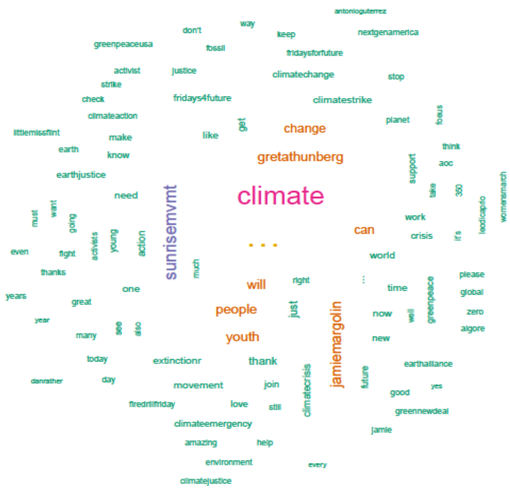

Fig. 2. #ThisIsZeroHour wordcloud.

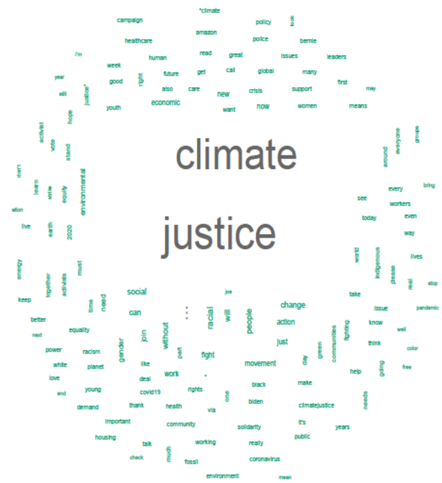

Fig. 3. #ClimateJustice wordcloud.

#ClimateJustice conveys 76.7% positive and 23.3% negative sentiments indicating that most people are happy with the movement. Emotions associated with the #ClimateJustice hashtag include disgust, surprise, sadness, anger, fear, joy, anticipation and trust in increasing order of polarity score. This implies that most people expressed joy, anticipation and trust in the #ClimateJustice action than the few who expressed disgust, surprise and sadness. The emotion scores for the #WeDontHaveTime hashtag are 1900, 3100, 1000, 3000, 2200, 1750, 1200, 4050, 4100 and 8150 for anger, anticipation, disgust, fear, joy, sadness, surprise, trust, negative and positive respectively as shown in Figure 7. #WeDontHaveTime hashtag conveys 66.7% positive sentiments and 33.3% negative sentiments indicating that most people are happy with the movement. The emotions associated with the #WeDontHaveTime include disgust, surprise, sadness, anger, joy, fear, anticipation and trust in increasing order of polarity score. Thus, most people expressed anticipation and trust in the #WeDontHaveTime action than the few who expressed disgust, surprise and sadness. We observed that a significant number of people expressed fear most especially with #WeDontHaveTime, followed by #ClimateJustice and #ThisIsZeroHour. However, to achieve meaningful progress in real-life, world nations and stakeholders should synergize efforts to limit greenhouse gases, ending deforestation and coal consumption, and expand financial aid for minimizing and adapting to climate change.

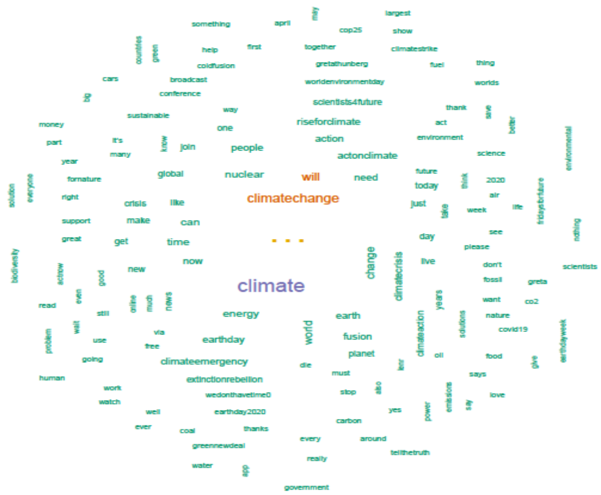

Fig. 4. Wordcloud of #WeDontHaveTime hashtag.

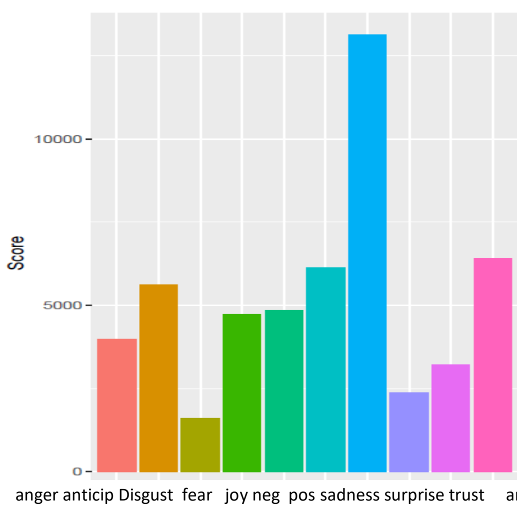

Fig. 5. Sentiment and emotion analysis of the #ThisIsZeroHour.

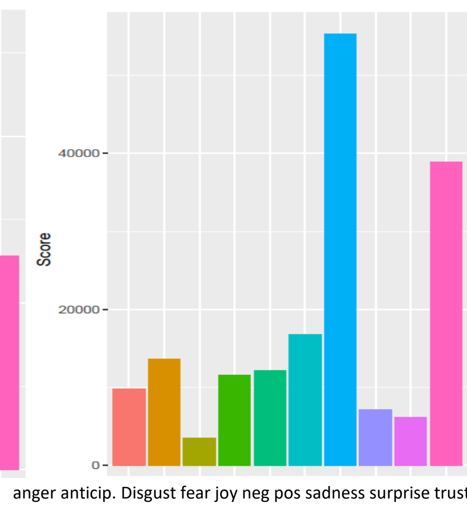

Fig. 6. Sentiment and emotion analysis of the #ClimateJustice.

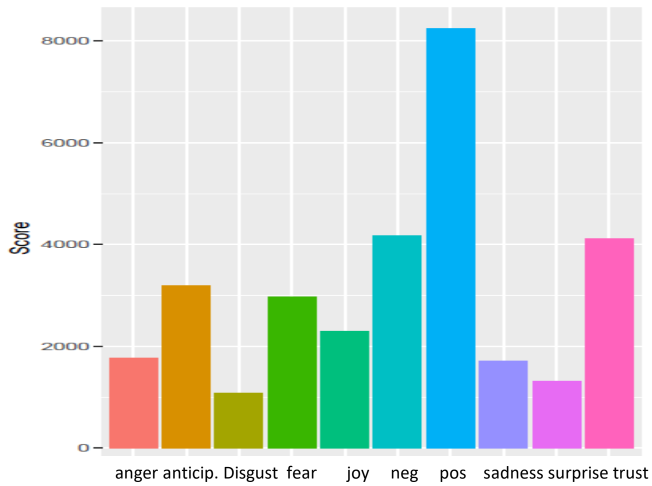

Fig. 7. Sentiment and emotion analysis of #WeDontHaveTime.

5 Conclusion

The primary focus of the global youth climate action has been towards changing and influencing the behavior and priority of government as well as to create profound awareness globally on the challenges of climate change. In this paper, sentiment and emotion analysis of the global youth climate action was conducted under the three prominent hashtags consisting of the #WeDontHaveTime, #ThisIsZeroHour and #ClimateJustice on Twitter in order to gain insight into the public emotions and sentiments that surround the climate change action. An insight obtained from the #ThisIsZeroHour wordcloud include the prominent roles of few activists at drawing attention towards the need to address climate change. In all the evaluations, most people expressed joy, anticipation and trust emotions in the #ThisIsZeroHour and #ClimateJustice action than the few who expressed disgust, sadness and surprise. The #ClimateJustice conveys the most positive sentiments, followed by #ThisIsZeroHour and the #WeDontHaveTime in that order. In all, a significant number of people express fear in the climate action. This corroborates the report of [13] that fear and hope are the major reasons for the youth climate change movement. Future work may consider processing spatio-temporal metadata and potential sarcasm, irony and satire contained in tweets while conducting emotion and sentiment analysis of similar events.

References

1. IPCC Climate Change Synthesis Report. Contribution of working groups I, II and III to. the fifth assessment report of the Intergovernmental on Climate Change. Climate change 2014 synthesis report. IPCC, Geneva, Switzerland. (2014).
2. Fagbola, T.M. & Thakur, S.C., Lexicon-based Bot-aware public emotion mining and sentiment analysis of the Nigerian 2019 presidential election on Twitter. *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 10, pp. 329-336. (2019).
3. Lineman, M., Do, Y., Kim, J.Y. & Joo, G.J., Talking about climate change and global warming. PLoS ONE vol. 10, no. 9: e0138996. (2015). https://doi.org/10.1371/journal.pone.0138996.
4. Biraj, D., Sathish, A. P. K., & Zhenlong, L., Topic modeling and sentiment analysis of global climate change tweets. Social Network Analysis and Mining, vol 9, no. 24, (2019).
5. Jost, F., Dale, A., & Schwebel, S., How positive is “change” in climate change? A sentiment analysis. Environmental Science and Policy, vol. 96, no. 2019, pp 27–36, (2019).
6. Abbar, S., Zanouda, T., Berti-Equille, L., & Borge-Holthoefer, J., Using Twitter to understand public interest in climate change: The case of Qatar,” The Workshop Program of the Tenth International Conference of Association for the Advancement of Artificial Intelligence’s on Web and social media (AAAI-16), Cologne, Germany. Technical Report WS-16-20, May 17, (2016).
7. Pathak, N., Henry, M. & Volkova, S., Understanding social media’s take on climate change through large-scale analysis of targeted opinions and emotions,” Association for the Advancement of Artificial Intelligence (www.aaai.org), (2017).
8. Xiaoran, A., Auroop, G.R., Steven, B.S., Hunter, A. M. & Dy, J. G., Tracking climate change opinions from Twitter data. Workshop on Data Science for Social Good held in conjunction with KDD2014, August 24, 2014, New York City, USA. (2014).
9. Ojala, M. & Lakew, Y. Young people and climate change communication. Oxford Research Encyclopedia of Climate Science, Climate Change Communication, Online Publication, (2017). doi: 10.1093/acrefore/9780190228620.013.408.
10. Chavoshi, N., Hamooni, H. & Mueen, A., DeBot: Twitter Bot Detection via Warped Correlation,” in ICDM, pp. 817-822, (2016).
11. Plutchik, R., The psychology and biology of emotion, New York: Harper Collins, (1994).
12. https://www.rollingstone.com/results/#?q=zero-hour-%20youth-climate-activist-jamie-margolin-video-979679. Accessed on 25 September, (2021).
13. https://www.latimes.com/environment/story/2019-09-19/climate-change-youth-activism. Accessed on 25 September, (2021).