

Lexicon-based Sentiment Analysis and Emotion Classification of the Global Youth Climate Protest on Twitter

Temitayo Matthew Fagbola¹ Abdultaofeek Abayomi² Murimo Bethel Mutanga³ and Vikash Jugoo⁴

¹ Department of Computer Science, Federal University Oye-Ekiti,
PMB 373, Ekiti State, Nigeria
temitayo.fagbola@gmail.com

^{2,3,4} Department of Information and Communication Technology, Mangosuthu University of Technology,
P.O.Box 12363, Jacobs, 4026 Durban, South Africa
taofeekab@yahoo.com

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, stakeholders and policy makers should consider the sentiments and emotions expressed by people and incorporate solutions in their various programmes toward addressing the climate change challenges especially as it affects nature and the ecosystem.

Keywords: Climate_Change, ClimateJustice, Climate-Protest, Emotion_Classification, Global Youth, Sentiment-Classification, ThisIsZeroHour, WeDontHaveTime

I. Introduction

Climate change and global warming 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. These are either in the form of extreme weather conditions such as drought, flood, increase in temperature or wildfires and declining crop yields. Furthermore, 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. These negative impacts traceable to climate change include low crop yields, ocean acidification and, increased frequency and intensity of daily temperature extremes [1]. In order to make nations more serious about climate change and move from business-as-usual situations to transformative change in various development paths, a lot of awareness campaigns and protests have taken place across the globe urging the nations to act now towards a safe nature, safe earth and a safe future which are important for various nature and biologically-inspired research studies. Some of these campaigns and protests are organised and undertaken by the youth as they constitute about 2.2 billion of the world's population. As a result of their challenging of government policies, 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.

The advent of the internet and social media networks have tremendously affected how people communicate and express opinions on global issues. It is therefore imperative to examine the awareness and knowledge of the people on climate change and its impacts as well as the sentiments that are formed around this global problem in order to be able to address a lot of concerns while also inspiring a collective success at all fronts in this regard. We leveraged on Twitter social media platform to harvest sentiments and emotions of users about climate change and conducted sentiment and emotion analysis so that public disposition towards the phenomenon can be measured.

With traditional sentiment analysis, a comprehensive interpretation of the contextual meanings associated with a text is limited. However, emotion mining approaches can be applied to gain more in-depth understanding of the user's perspective of a text. Emotion mining extracts 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 are the public sentiments and emotions regarding climate change actions globally? The rest of the paper is organized as follows: Section 2 presents the related works regarding climate change and sentiment analysis. Section III presents the materials and method while Section IV presents the results and discussion and in Section V the conclusion is presented.

II. 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 and knowledge of the terms climate change and global warming by assessing the relative search volume patterns (RSV) for the terms. Using the Twitter API, a total of 21,182 and 26,462 tweets referencing the terms climate change and global warming or their acronyms were collected between the period October 12 to December 12, 2013. Sentiment analysis and opinion mining was performed on the analysed dataset using the Semantria® software while the differences in RSV's for the terms, for the investigation period were identified using a paired t-test. Similarly, the Pettitt and Mann-Kendall tests were used to identify changes in distribution, averages and the presence of trends within the RSVs on weekly basis. The sentiments connected to the terms in social media networks was examined and the authors established a relationship between the awareness of information and the quantity of publicity created for climate change and global warming. The increase in awareness was triggered by increase in publicity which could either be positive or negative for the terms as the original context in which the terms were constructed is directly related to the emotive connections attached to them. Therefore, the knowledge and awareness on climate change and global warming is based on the exposure enjoyed by the terms in the (social) media.

In another study [4], Twitter data was used to evaluate public opinion on climate change. Geotagged tweets containing certain keywords such as carbon dioxide, fossil fuel, carbon footprint, emissions etc. that are related to climate change was analysed by using sentiment analysis and topic modelling text mining techniques. The dataset is made up of 390,016 tweets harvested globally for the period July 1, 2016 -February, 2018 using the Twitter Stream API. For the sentiment analysis, the overall feelings and attitudes of the public on climate change as expressed in the dataset was determined by applying the Valence Aware Dictionary and sEntiment Reasoner (VADER) while the various topics being discussed were deduced by

using the Latent Dirichlet Allocation (LDA) topic modeling technique. The types of climate change discussions in different countries over a period of time were analysed using the VADER and LDA techniques. The study found that varied topics on climate change are being discussed, though some are more popular than the others while the overall discussions on the issue of climate change is negative particularly in respect of tweeting relating to extreme weather conditions, government and institutional policy issues.

In order to assess the success recorded in the path of climate change actions using its drivers and barriers, the authors in [5] explain the results of a non-computational sentiment analysis of the word "change" on the term development path change, local climate innovation and climate change. In determining the actual occurrence of changes in each development path, sentiment analysis procedure was carried out to confront the lack of agreement among policy makers and climate change scholars. As a result of 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. A research project, tagged MC3 Meeting the Climate Change Challenge was conducted by the researchers over a six- year period to examine climate innovation in eleven Canadian local municipalities. An open-ended, semi-structured interview protocol that drew upon three theoretical lenses including multi-level perspectives, social practice theory, and social-ecological systems from across British Columbia, Canada was used. The positive changes and sentiments were connected to local government leadership and staff quality which exist in the institutions. This is to confirm that their efforts are yielding positive results. On the other hand, significant negative sentiments are connected to the magnitude and swiftness of change while the local government staff consider behavioural change as a critical barrier.

Abbar et al. 2017 [6] examined and analysed climate change tweets in Qatar, Arabian Peninsula in order to determine the public awareness of this global challenge and the different topics including politics, economy, energy, disasters and sandstorm etc. that are associated with it. The authors collected users' tweets and applied a taxonomy of climate change topics to extract discussions relating to this global issue. The various topics discussed were also investigated and the public gravitation towards them examined. Air quality trends was focused on, especially as the country is designated as the highest per-capital emitters of CO₂. A total of 109.6 million tweets was harvested from 98,066 different users for the period July 2014 to July 2016, using the Twitter API while daily weather conditions data consisting of temperature, humidity, wind speed, rain, snow, and visibility were also collected. After applying the content and correlation analyses to the datasets, the study was able to classify tweets into various topics covering a wide range of issues including air quality, politics and economy. Some of the findings include people's interests are driven by widely covered events that have direct link to their daily lives; and the number of users engaging in climate change discussion is not necessarily in an increasing trend.

In the research study reported by [7], a large collection of Tweets relating to the social conversation on global climate change during the 2015's 21st Conference of the Parties (COP21) of the United Nations on Climate Change, Paris was analysed based on emotions, opinions and demographics dynamics over time and location, in order to discover how the various users discuss, behave and form opinions on meaningful topics on the social media. Over 4.5 million English language tweets relating to the COP21 event, terms, hashtags and user accounts were processed. The tweets discourse were categorised into nine groups which include energy, weather, economy, agriculture/forestry, water, security, climate denial, air issues and animals. This is to enhance the analysis of targeted opinions on specific topics that are relevant to climate change. The tweets were also annotated with sentiments and emoticons by using publicly available affect models such that the sentiment class are positive, negative and neutral while joy, fear, sadness, surprise, anger and surprise belongs to the emotions category. 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.

An understanding of whether Twitter data mining can be used to supplement or complement insights into climate change knowledge and perceptions was studied by [8]. Text mining, time series and sentiment analysis techniques were applied to determine whether upon exposure to climate related hazards such as wildfire, hurricane, flood, drought etc., if such perceptions be changed over time. After pre-processing, a total of 494097 tweets relating to climate change during the period October, 3rd 2013 and December, 12th 2013 were harvested and analysed. For sentiment analysis, the authors manually labelled the tweets and classified them into subjective and objective classes. The subjective class expresses emotions or opinions about Climate Change by the user and this class is further partitioned into positive and negative groups. The objective class contains tweets of news or articles about climate change and this is partitioned into positive for such tweets that expresses concerns or actions to be taken to curb the global issue; and negative for other tweets that doesn't believe in climate change. The Naïve Bayes and support vector machines classification algorithms were applied on the labelled dataset for the sentiment text classification task. The Naïve Bayes classifier achieved the best performance of 76.54% accuracy, F1 score of 0.8057; and 76.77% accuracy, F1 score 0.8278 respectively for both the subjective and polarity classification tasks.

Ojala and Lakew [9] focused on the youth as a crucial group to mobilise and integrate into the efforts of curbing climate change consequences. The authors leveraged on the fact that the world's population is young and are societies' leaders in the future, opined that this group will be responsible for handling future negative effects of the global menace. Research was conducted bothering on climate change communication and the young people in order to enhance the comprehension of how the young people relates to issues of

global climate change and communicate in various context. This is aimed at promoting knowledge, a sense of efficacy and participation concerning climate change. The study examined how young people are influenced by communications on climate change and do use the social media to interact on the issue while also investigating their coping with the negative emotions triggered by knowledge and awareness about climate change. The authors found that the art-based approaches are vital ways to communicate with the young people on the climate change issue. The young people's knowledge and awareness about climate change is also easily increased compared to making them more engaged and taking responsibilities around the issue. Lastly, there are identified variations across media channels and audiences on the role that the mass media play in young people's climate change awareness and engagement.

Gabrielle and Irina [14] investigated the relationship between emotional response, behaviours, intentions and attitudes relating to climate change. Patronella [15] conducted a sentiment analysis from over 13,000 climate change articles published between 2010 and 2020 by 11 key newspapers in the United States of America, including The Washington Post, The New York Times, The wall Street Journal, Boston Globe, Tampa Bay Times, USA Today, Star Tribune, St Louis Post Dispatch, Pittsburgh Post-Gazette, Philadelphia Inquirer, and Atlanta Journal-Constitution. Factiva was used to build the corpus based on search key terms like "climate change" or "global warming". The author combined both automated approaches to emotion detection and manual annotation of emotions to identify the key emotions contained in the articles. The valence and emotions inherent in the articles were also analyzed by using the Tidy Text NRC Word-Emotion Association Lexicon (EmoLex) and the Lexicoder Sentiment dictionaries. These dictionaries contain keywords and their associated emotions and valence, respectively. The top emotions obtained include trust which is around 27 percent, anticipation (17 percent) and fear which accounts for 14 percent of the total emotions.

Tishya, Ashwani and Vibha [16] investigated how the intensity of emotion and public opinions from climate change tweets could be quantified. A framework based on Bi-directional Long Short Term Memory (Bi-LSTM) and a lexicon emotion analyzer was used to evaluate the intensity of public opinion and emotions from climate change tweets. Yasin and Seher [17] investigated the sentiments contained in a corpus of tweets in Turkish language relating to climate change and global warming. The authors adopted both linear and probabilistic classification over 30,000 Turkish tweets which were randomly selected for emotion detection purposes. A N-gram technique was used for tokenization while Naïve Bayes, K-Nearest Neighbour and Support Vector Machine were used for classification.

Zhen [18] employed a multinomial Naïve Bayes algorithm to conduct opinion mining model training and sentiment analysis of 100,000 climate change-related tweets of Twitter users. Top three (3) climate change hashtags of interests employed in the study to fetch the relevant tweets were discovered through the best-hashtags.com. These top hashtags include the

#climatechange, #globalwarming and #environment. Preprocessing of the tweets was achieved using the SKLEARN and NLTK libraries in Python through the tokenization, stop words removal, morphological normalization and collocation processes. The results obtained indicate that more people expressed negative sentiment regarding climate change rather than the few who expressed optimistic positive sentiment towards climate change. This result has established a strong indication that climate change stakeholders should pay immediate attention to fighting climate change.

III. 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 2. 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 analytical engine and output visualization.

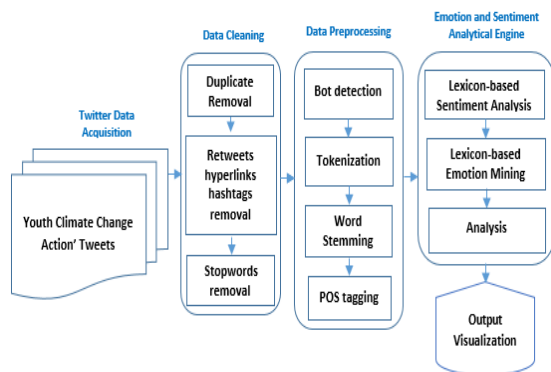


Figure 1. Our lexicon-based sentiment and emotion analyzer [2].

A. 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. In Table 1, the number of tweets and retweets associated with the hashtags is summarized and presented.

Table 1. Youth climate Action Hashtags and associated tweets.

Hashtags	No. of Tweets	No. of Retweets
ThisIsZeroHour	10,120	41,077
ClimateJustice	19,449	144,484
WeDontHaveTime	6,706	36,172

As shown in Table 1, 10120 tweets and 41077 retweets were obtained using the #ThisIsZeroHour. Similarly, 19449 tweets and 144484 retweets are connected with #ClimateJustice. The #WeDontHaveTime produced 6706 tweets and 36172

retweets. Samples of tweets from #ThisIsZeroHour, #ClimateJustice and #WeDontHaveTime are presented in Figures 2, 3 and 4.

"I want a secure future for my generation and the generations after me."

"It also showed a key ignorance that has been exploited by the gas industry: The public doesn't understand that the gas is essentially the same as methane, a pollutant."

"We need influencers to promote upcycling, 2nd hand clothing, and circular economy"

"The main cultural disorder that produces socioclimatic injustice originates in the CULT OF THE EGO; that derives in multiple racisms, classisms, ageisms, nationalisms and suprematisms."

Figure 2. Sample tweets from #ThisIsZeroHour hashtag.

"Everything that destroys the planet and kills nature kills us"

"Without resistance Africa will stay in the pot like the proverbial frog in the pan, barely noticing the rising heat among so many other survival considerations, until she is cooked in the cauldron. The more reason Africans should amplify their voice in calling for."

"Every time we purchase a product, any product, there is a natural resource or human resource somewhere on the planet that becomes poorer. Our economic system cannot continue like this"

"On the beach and in the sea, Animals don't leave Trash; Humans do. Please behave like animals."

Figure 3. Sample tweets from #ClimateJustice hashtag.

"Unprecedented forest fires 🌳 are happening in #Argentina right now. 800 000 m² has been destroyed so far!"

"The new #IPCC report will be published on Monday and will stress urgency. If rapid action is taken on cutting greenhouse gases and spending on helping people adapt to climate change is ramped up, then the worst risks can be avoided."

"I dream of a sane world where science and rationality rule the day. Sooner the better. That means no trolling and wacko conspiracy theories. Who's with me?"

"Scientists raise alarm over 'dangerously fast' growth in atmospheric methane"

"What will solve the climate crisis is that we need people in communication. We need behaviour scientists and we need to understand why people are so passive"

Figure 4. Sample tweets from #WeDontHaveTime hashtag.

B. Data cleaning and preprocessing

A lot of noise is often associated with social media data including tweets. Hence, data cleaning becomes highly imperative with the objective of detecting and removing the noise contained in the data. With this, more relevant and reliable data was obtained for further analysis. Twitter handles, stopwords, hyperlinks, special and repeated characters in tweets, duplicate hashtags and tweets, empty spaces and retweet entities were all removed using the appropriate functions defined in the documentation in 'R'.

Furthermore, Data preprocessing puts the data in a ready-to-use state for a reliable, quality-assured and convenient analysis. This step involves bots detection and removal, tokenization, stemming and part-of-speech tagging. Bots contained in the twitter data collected were detected using the Botometer [10]. With a bot score of 0.6, 1063 tweets

The emotion and sentiment 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 sentiment and positive sentiment respectively as displayed in Figure 10. #WeDontHaveTime hashtag conveys 66.7% positive sentiments of the overall sentiment distribution 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.

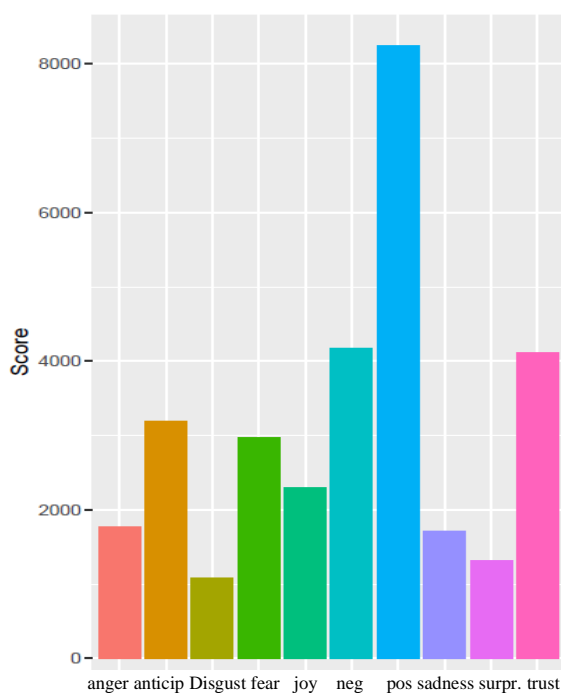


Figure 10. Sentiment and Emotion Analysis of the WeDontHaveTime hashtag.

V. 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. This is important in view of the effects of climate change on nature and ecosystem consisting of the fauna and flora on which nature and biologically inspired research studies are conducted. 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. The 10120, 19449 and 6706 tweets posted between 01 September 2019 and 28 Jun, 2020 were obtained for the #ThisIsZeroHour, #ClimateJustice

and the #WeDontHaveTime hashtags respectively, via the Twitter's streaming API. The NRC word emotion association lexicon dictionary and the Syuzhet library functions were used to detect and estimate polarity of sentiments of positive and negative, and Plutchik's wheel of emotions which consist of anger, anticipation, disgust, fear, joy, sadness, surprise, and trust expressed in the tweets that originated from the global youth climate action. An insight obtained from the #ThisIsZeroHour wordcloud include the prominent roles of activist in 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 works would be directed towards incorporating spatio-temporal metadata from tweets in emotion and sentiment analysis of related events. In a similar manner, linguistic devices including sarcasm, irony and satire which usually influence the outcome of sentiment analysis in text will be addressed in future works.

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] T.M. Fagbola & S.C. Thakur, 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] M. Lineman, Y. Do, J.Y. Kim, & G.J. Joo, Talking about climate change and global warming. *PLoS ONE* vol. 10, no. 9: e0138996. (2015). <https://doi.org/10.1371/journal.pone.0138996>.
- [4] D. Biraj, A.P.K. Sathish & L. Zhenlong, Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, vol 9, no. 24, (2019).
- [5] F. Jost, A. Dale, & S. Schwebel, How positive is "change" in climate change? A sentiment analysis. *Environmental Science and Policy*, vol. 96, no. 2019, pp 27-36, (2019).
- [6] S. Abbar, T. Zanouda, L. Berti-Equille & J. Borge-Holthoefer, 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] N. Pathak, M. Henry & S. Volkova, 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] A. Xiaoran, G.R., Auroop, B.S. Steven, A.M. Hunter & J. G. Dy, 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] M. Ojala & Y. Lakew, 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] N. Chavoshi, Hamooni, H. & Mueen, A., DeBot: Twitter Bot Detection via Warped Correlation,” in ICDM, pp. 817-822, (2016).
- [11] R. Plutchik, 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).
- [14] W. Gabrielle Wong-Parodi & F. Irina Feygina (2021): Engaging People on Climate Change: The Role of Emotional Responses, Environmental Communication, DOI: 10.1080/17524032.2020.1871051.
- [15] A. M. Patronella (2021). "Covering Climate Change: A Sentiment Analysis of Major Newspaper Articles from 2010 - 2020" *Inquiries Journal*, 13(09).
- [16] T. Tishya Thukral, A. Ashwani Varshney and V. Vibha Gaur. Intensity quantification of public opinion and emotion analysis on climate change. International Journal of Advanced Technology and Engineering Exploration, Vol 8(83), 2021, pp. 1351-1366.
- [17] Yasin Kirelli and Seher Arslankaya. Sentiment Analysis of Shared Tweets on Global Warming on Twitter with Data Mining Methods: A Case Study on Turkish Language. Hindawi, Computational Intelligence and Neuroscience, PP. 1-9, Volume 2020, Article ID 1904172, <https://doi.org/10.1155/2020/1904172>.
- [18] Zhen Wang. Sentiment Analysis of Climate Change Using Twitter API and Machine Learning. Bachelor's Thesis, Engineering, Information and Communication Technology, Turku University of Applied Sciences, 2020, pp. 1-39.