University of Exeter coursework header sheet

ECMM451 Data Science Research Project (A, TRM123 2022/3)

240829

1086732

720079564

Coursework: Project Proposal

Submission Deadline: Sat 6th May 2023 12:00

Personal tutor: Professor Jonathan Fieldsend

Marker name: N/A

Word count: 0

By submitting coursework you declare that you understand and consent to the University policies regarding plagiarism and mitigation (these can be seen online at www.exeter.ac.uk/plagiarism, and www.exeter.ac.uk/mitigation respectively), and that you have read your school's rules for submission of written coursework, for example rules on maximum and minimum number of words. Indicative/first marks are provisional only.

Trust and trolling of weather forecasters: An analysis of Twitter conversations related to climate change during UK heatwaves in 2020 and 2022

Asfee Alam
Department of Computer Science
University of Exeter
Exeter, UK
aa1308@exeter.ac.uk

***Abstract*—This project proposes to analyze Twitter conversations related to climate change during UK heatwaves in 2020 and 2022, with a focus on identifying patterns of abusive behavior directed towards weather forecasters. The proposed methodology involves gathering relevant tweets and analyzing the tweet content using natural language processing (NLP) techniques to determine the type and degree of abuse as well as language usage patterns that are indicative of abusive behavior. The study will apply topic modeling techniques to identify the key topics or themes that are being discussed in tweets. In order to spot any shifts or trends, the study will also analyze the patterns of abuse in 2020 and 2022. The findings of this study will provide valuable insights into the attitudes and behaviors of Twitter users towards climate change and weather forecasters and help inform the development of interventions to address online abuse in the context of climate change discourse. Ultimately, this study aims to contribute to a more informed and respectful public discourse around climate change and related issues. The research question is: What are the attitudes and behaviors of Twitter users towards climate change and weather forecasters during heatwaves, and how can automated methods be used to provide insights into these phenomena?**

***Index Terms*—Sentiment Analysis, Abuse Detection, NLP Analysis**

I. Introduction

Climate change has become a topic of global concern in recent times, with increasing temperatures being one of its most significant impacts. As a result of climate change, heatwaves—prolonged periods of very hot weather—have increased in frequency and intensity. These heatwaves are of interest to both researchers and policymakers because of the serious effects they have on human health, the environment, and the economy. Recent heatwaves in the UK have raised a lot of discussion and concern about how climate change may be contributing to their occurrence.

One of the key obstacles to effective action on climate change is a lack of public awareness and understanding of the issue. In recent years, social media platforms like Twitter have become important tools for disseminating information about heatwaves and climate change. However, they have also become breeding grounds for online abuse and trolling. Weather forecasters who report on heatwaves and climate change on Twitter have been subjected to unprecedented levels of trolling and abuse from users who deny the existence of climate change or dispute the link between heatwaves and climate change. This trolling can have serious implications for the mental health and well-being of the forecasters and can also undermine public trust in scientific information related to climate change.

According to a report by the BBC, weather forecasters reported unprecedented levels of trolling in response to Twitter posts comparing the UK heatwave in 2022 to climate change[3]. This highlights the need for research that investigates the nature and extent of this trolling and identifies factors that contribute to it. This problem highlights the difficulties of informing the public with accurate and pertinent information on climate change. It also raises questions about the impact of social media on public perceptions of climate change and the role of weather forecasters in communicating climate change to the public. Therefore, in this proposal, we propose to analyze tweets replying to or mentioning specific users during the periods of UK heatwaves in both 2020 and 2022 to identify the type and level of abuse, the use of language in these posts, and any underlying factors that can be determined. This research will contribute to our understanding of the impact of climate change and heatwaves on online behavior and will help identify ways to mitigate the effects of online abuse on climate change communication.

II. Research Context

According to research on the subject, climate change is a complex and difficult topic to communicate to people in general[13]. It is a problem that has a big impact on weather patterns, particularly how often and how intense heatwaves are. It may have been challenging to explain and communicate in a way that was relatable to the scientific nature of climate change and its potential effects. Effective communication about climate change has been emphasized by researchers, and social media sites like Twitter have a part to play in providing

a platform for public engagement[7]. Studies, however, have also drawn attention to the role that social media plays in the dissemination of false information and the potential for online abuse toward those involved in climate change communication[16].

Weather forecasters have recently taken a prominent role in highlighting the severity and effects of climate change. However, in reaction to tweets discussing the connection between heatwaves and climate change, they have been subjected to unprecedented levels of trolling on Twitter[3]. This situation calls into question the usefulness of social media as a medium for communicating about climate change as well as the impact of abuse and trolling on weather forecasts. Recent reports and studies have emphasized the necessity to address the abuse directed towards weather forecasters and to understand the underlying elements contributing to this phenomenon.

Studies on social media abuse and trolling have identified the role of anonymity, perceived anonymity, and the formation of online communities as factors contributing to the prevalence of such behavior[14]. Furthermore, research on the use of language on social media has shown that certain patterns of language use can be indicative of abusive behavior[12].

Sentiment analysis, which is the act of determining the polarity of a text, such as positive, negative, or neutral, can be used to analyze social media data. In the context of climate change, sentiment analysis can be used to analyze public attitudes toward climate change and to evaluate the effectiveness of climate change communication strategies. Sentiment analysis has previously been used to discover and monitor public sentiment regarding climate change on Twitter[4].

Another approach is using an abuse detection algorithm, a machine learning algorithm designed to automatically identify abusive content in social media. Previous studies have used abuse detection algorithms to analyze social media data related to topics such as politics, sports, and entertainment[5]. In the context of climate change, an abuse detection algorithm can be used to identify and measure the prevalence of abusive content directed toward climate change communicators and forecasters.

Overall, the literature suggests that social media data relating to climate change and weather forecasting can be analyzed using a mix of sentiment analysis, topic modeling, and an abuse detection algorithm. The effectiveness of climate change communication techniques, public attitudes towards the issue, and the frequency of offensive content directed toward climate change communicators and forecasters can all be significantly impacted by these methodologies.

III. Aims & Objectives

The aim of this study is to investigate Twitter conversations related to climate change during heatwaves and to develop a framework to better understand the attitudes and behaviors of Twitter users towards climate change and weather forecasters.

To achieve this aim, the following objectives will be pursued:

- To collect and pre-process a dataset of Twitter posts related to climate change and heatwaves.
- To perform sentiment analysis on the collected tweets to understand the overall sentiment of Twitter users towards climate change and weather forecasters during heatwaves.
- To use topic modeling techniques such as Latent Dirichlet Allocation (LDA) to identify the key topics being discussed on Twitter during heatwaves related to climate change and weather forecasters.
- To develop an abuse detection algorithm using Natural Language Processing (NLP) techniques to identify abusive language directed toward weather forecasters or climate change communicators.
- To analyze the data to identify any patterns or correlations between sentiment, topics of discussion, and abusive tweets.

IV. Data & Resources

The dataset selected for this project consists of tweets collected during the heatwaves that occurred in the UK between July and August 2020 and 2022. This time frame is chosen because both years had heatwaves that led to significant impacts on public health, infrastructure, and the environment. Also, in both these years, weather forecasters received massive amounts of abuse and trolling tweets. Thus, this timeline will be perfect to get a grasp of how the public reacted to the heatwave and their reaction to meteorologists' tweets. The dataset is collected from the Twitter API which allows access to the platform's real-time and historical data. Some keywords and hashtags that are used to extract the data are, #heatwave, #heatwaveuk, #heat, #summer, and #hot. Also, tweets from a few weathers forecaster's accounts like @MetMattTaylor, @alexdeakin, @metoffice, @Schafernaker, and @DrLizBentley are also taken as they had a lot of tweets which had received a lot of trolls in the replies. The proposed study will rely on open-source software tools such as Python, Jupyter Notebooks, and relevant libraries for data processing, analysis, and visualization. Overall, this study will rely heavily on the use of Twitter data to gain insights into the public's attitudes towards climate change and weather forecasters during heatwaves.

V. Methods & Experimental Design

The proposed methodology for this study involves the following steps:

1) *Data Pre-processing and Analysis*: The data will be pre-processed by using natural language processing (NLP) techniques. Python programming language and the NLTK library will be used for pre-processing tasks such as tokenization, stop-word removal, and stemming. The pre-processed data will be used for sentiment analysis, topic modeling, and abuse detection.
2) *Sentiment Analysis*: For sentiment analysis, we will use the VADER (Valence Aware Dictionary and Sentiment Reasoner) tool, which is a lexicon and rule-

based sentiment analysis tool. Each tweet will receive a sentiment score from the tool, which we will use to evaluate Twitter's general opinion on climate change and heatwaves. For this project, we will be using sentiment analysis at the phrase level to check the negative and positive sentiment along with its polarity.

3) *Topic Modeling*: Latent Dirichlet Allocation (LDA), which is a widely used unsupervised machine learning technique for identifying topics in text data will be used. LDA will help us identify the most common topics discussed on Twitter related to climate change and heatwaves. We will use the topic modeling results to identify the key themes and issues that are driving the Twitter conversation around climate change and heatwaves.
4) *Abuse Detection*: To find tweets that contain offensive language addressed at weather forecasts, an abuse detection algorithm will be created. In order to do this, NLP techniques will be used to recognize language patterns that point to abusive behaviors.
5) *Model Evaluation*: To evaluate the performance of our models, we will use a variety of metrics, including accuracy, precision, recall, and F1 score.
6) *Data Visualisation*: The results of the analysis will be presented in the form of data visualizations, such as graphs, word clouds, and charts. This will help to provide a clearer understanding of the patterns and trends in the data.

VI. Data Governance & Ethics

Several ethical and legal issues are brought up by the usage of social media data in research. This project will strictly follow data governance and ethical rules throughout the whole research process in order to mitigate these worries. All the data for this study will be gathered via Twitter's free and publicly accessible public APIs. The data will be de-identified and aggregated in order to protect user confidentiality.

We will also make sure to comply with all relevant data protection requirements, such as the GDPR and UK data protection laws. Furthermore, we'll make sure that none of the methods we use to analyze the data stigmatize or discriminate against any particular people or group.

The security and confidentiality of all the data used in the study will also be maintained. Data will be kept on password-protected devices, and only authorized people will be able to access it. To safeguard the privacy of specific users, any published or shared data will be de-identified and aggregated. After the project is over, the data will be erased out of concern for security and privacy.

VII. Project Plan

The proposed project will be conducted over a period of 28 weeks starting from 14th Feb 2023 to 12th August 2023. The following is a brief timeline of the major tasks involved in the project:

- *Literature Overview (28 Weeks)*: It is necessary to update and cite a literature review on a regular basis. Due to this precise reason, several papers, journals, books, etc. have been researched since the project's beginning, and the ones that are the most pertinent to it have been added to the references. These references will continue to grow as the project nears its conclusion.

- *Data Pre-processing (18 May, 23 - 1 June, 23)*: Pre-process data by removing stop words, stemming, and tokenizing the text. Perform exploratory data analysis to gain insights about the data.

- *Sentiment Analysis (2 June, 23 - 16 June, 23)*: Conduct sentiment analysis on the pre-processed Twitter data using Natural Language Processing (NLP) techniques such as VADER. Analyze sentiment distribution to understand the overall sentiment of tweets related to heatwaves and climate change.

- *Topic Modeling (17 June, 23 - 1 July, 23)*: Apply Latent Dirichlet Allocation (LDA) to identify the underlying topics in the preprocessed tweets. Analyze topic distribution and coherence to understand the key themes discussed on Twitter related to heatwaves and climate change.

- *Abuse Detection (2 July, 23 - 16 July, 23)*: Develop an abuse detection algorithm using supervised learning techniques based on language patterns associated with abusive behavior. Apply the algorithm to identify tweets containing abusive language towards weather forecasters and climate change activists.

- *Experimentation (17 July, 23 - 24 July, 23)*: Conduct experiments to investigate the relationship between sentiment, topics, and abusive language in tweets related to heatwaves and climate change. Analyze the results and conclude the attitudes and behaviors of Twitter users towards climate change and weather forecasters.

- *Presentation and Research Paper Writing (25 July, 23 - 15 August, 23)*: Write a research paper based on the findings of the experiments. Revise and edit the paper based on feedback from peers and supervisor. Prepare and deliver a final presentation of the project.

VIII. Risk Assessment

Several risks could prohibit me from achieving the specified objectives during the course of this research. The project's use of incorrect data raises some issues. A lot of noisy data and spam messages are common on Twitter. Several data cleaning procedures will be used, such as removing spam

tweets, eliminating irrelevant phrases, and removing retweets, to reduce the danger of incorrect data. In order to guarantee that the information gathered is representative of the full dataset, random sampling will also be performed.

The second issue is the inability to train the two models within the time allotted for each model. Some models require more time to train and perfect, but because this project has a set finish date, both models must be correctly trained, fine-tuned, and tested before the research paper is completed. Pre-trained models will be used as a starting point to reduce training time, and the models will be adjusted on the gathered data to decrease the risk. To expedite the training process, parallel computing techniques will also be used.

The possibility of technical problems or constraints represents another risk. This might happen as a result of things like hardware or software malfunctions, restrictions on the amount of memory available, or problems with the natural language processing methods being employed. I'll use the proper hardware and software tools, as well as carry out thorough testing and validation of the employed natural language processing approaches, to reduce this danger. Overall, I will keep a close eye on the project's development and take the necessary steps to reduce any potential hazards.

IX. Conclusion

In conclusion, this proposal provides a research project that aims to investigate Twitter users' attitudes and behaviors towards climate change and weather forecasts during heatwaves in 2020 and 2022, with an emphasis on finding patterns of abusive behavior targeted against meteorologists. This study will offer significant insights into the public debate on climate change and how it is transmitted on social media platforms by utilizing machine learning techniques like sentiment analysis, topic modeling, and an abuse detection algorithm. The study will add to the expanding corpus of research on social media analysis, online abuse, and trolling as it relates to climate change communication. The results of this study will have applications for lawmakers, social media platform designers, weather forecasters, and anyone who communicates about climate change. Overall, this research has the potential to advance our knowledge of how climate change is communicated on social media and contribute to the development of more effective strategies for addressing this pressing issue.

X. References

1) Holm, Amalie & Doksæter Sivle, Anders & Jeuring, Jelmer & Svehagen, Mai-Linn. (2020). Climate Communication: TV meteorologists as Climate Communicators.
2) The Alan Turing Institute. (n.d.). Tracking abuse on Twitter against football players in the 2021-22 Premier League Season. [online] Available at: https://www.turing.ac.uk/research/publications/tracking-abuse-twitter-against-football-players-2021-22-premier-league-season.
3) UK heatwave: Weather forecasters report unprecedented trolling. (2022). BBC News. [online] 29 Jul. Available at: https://www.bbc.com/news/uk-62323048.
4) Sanford, M., Painter, J., Yasseri, T. and Lorimer, J. (2021). Controversy around climate change reports: a case study of Twitter responses to the 2019 IPCC report on land. Climatic Change, 167(3-4). doi:https://doi.org/10.1007/s10584-021-03182-1.
5) De Souza, A. and Abreu, D. (n.d.). Automatic offensive language detection from Twitter data using machine learning and feature selection of metadata. [online] Available at: https://shura.shu.ac.uk/26018/1/IEEE_WCCI_Artificial_Intelligence.
6) Cody, E.M., Reagan, A.J., Mitchell, L., Dodds, P.S. and Danforth, C.M. (2015). Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll. PLOS ONE, 10(8), p.e0136092. doi:https://doi.org/10.1371/journal.pone.0136092.
7) Brossard, D. and Nisbet, M.C. (2006). Deference to Scientific Authority Among a Low Information Public: Understanding U.S. Opinion on Agricultural Biotechnology. International Journal of Public Opinion Research, 19(1), pp.24–52. doi:https://doi.org/10.1093/ijpor/edl003.
8) Roser-Renouf, Connie & Maibach, Edward & Leiserowitz, Anthony. (2009). Global Warming's Six Americas: An Audience Segmentation Analysis (Invited). AGU Fall Meeting Abstracts.
9) Sinnenberg, L., Buttenheim, A.M., Padrez, K., Mancheno, C., Ungar, L. and Merchant, R.M. (2017). Twitter as a Tool for Health Research: A Systematic Review. American Journal of Public Health, 107(1), pp.e1–e8. doi:https://doi.org/10.2105/ajph.2016.303512.
10) Cook, J., Lewandowsky, S. and Ecker, U.K.H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. PLOS ONE, 12(5), p.e0175799. doi:https://doi.org/10.1371/journal.pone.0175799.
11) Capstick, S., Whitmarsh, L., Poortinga, W., Pidgeon, N. and Upham, P. (2014). International trends in public perceptions of climate change over the past quarter century. Wiley Interdisciplinary Reviews: Climate Change, 6(1), pp.35–61. doi:https://doi.org/10.1002/wcc.321.
12) Davidson, T., Warmsley, D., Macy, M. and Weber, I. (2017). Automated Hate Speech Detection and the Problem of Offensive Language. arXiv:1703.04009 [cs]. [online] Available at: https://arxiv.org/abs/1703.04009.
13) Moser, S. and Dilling, L. (2020). Communicating Climate Change: Closing the Science-Action Gap. [online] Available at: https://www.semanticscholar.org/paper/Communicating-Climate-Change
14) Suler, J. (2004). The online disinhibition effect. CyberPsychology & Behavior, 7(3), pp.321–326. doi:https://doi.org/10.1089/1094931041291295.

15) Fichman, P. and Rathi, M. (2023). Trolling CNN and Fox News on Facebook, Instagram, and Twitter. Journal of the Association for Information Science and Technology. doi:https://doi.org/10.1002/asi.24753.
16) Allison-Hope, D., Barrett, N., Coles, K., Darnton, H., Matthews, C. and Rohwer, M. (n.d.). Available at: https://www.fordfoundation.org/media/7376/bsr-15-07-22.pdf.
17) Blei, D., Ng, A. and Jordan, M. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, [online] 3, pp.993–1022. Available at: https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf.
18) Founta, A., Djouvas, C., Chatzakou, D., Leontiadis, I., Blackburn, J., Stringhini, G., Vakali, A., Sirivianos, M. and Kourtellis, N. (2018). Large Scale Crowdsourcing and Characterization of Twitter Abusive Behavior. Proceedings of the International AAAI Conference on Web and Social Media, 12(1). doi:https://doi.org/10.1609/icwsm.v12i1.14991.