*Disaster Scanner*

*Using Twitter automation to Recognize Natural Disasters and Quantify Emotional Impact*

|  |  |  |  |
| --- | --- | --- | --- |
| Hector Dondiego  hectordondiego@lewisu.edu | William Pulkownik  williamrpulkownik@lewisu.edu | Alec Schopler  alecpschopler@lewisu.edu | Tyler Starkus  tylerjstarkus@lewisu.edu |

***Abstract*— Natural Disasters can be unexpected and overwhelming experiences, and Twitter has become an important communication channel in times of emergency. In this paper we propose the construction of an automated Twitter tool whose function is to detect and identify possible disasters and classify them as genuine emergencies or false positives. We chose Twitter because of its robust developer API, and because tweets are widely-used, public, iterable, and tend to contain language rich in semantics. We intend to couple modern AI methods of natural language processing with established methods of social media automation to create a compound tool which may be of interest to Disaster Relief Agencies or media outlets. We hope one day this research may be used by relief agencies to quantify the emotional impact of disasters and aid in relief efforts.**

1. Introduction

In 2019, over 139 million people globally use Twitter every single day [7]. Within a given day, Twitter users send roughly 500 million tweets [7] discussing a vast array of topics such as entertainment, politics, and world news.With the amount of data that is generated from Twitter’s user base, significant events such as disasters can go unnoticed. Prior to the late 2010’s, working through this data within Twitter was a near impossible feat with the technology at the time. However, by using modern computing and artificial intelligence, it is possible to develop software that can accurately identify and categorize disasters within a reasonable timeframe by using tweets. As 22% of the US Adult population uses Twitter [7], the likelihood of new disasters being reported by Twitter users in the United States is high. The advantages of using such software to analyze Twitter tweets could provide a means for news stations to inform the public faster, relief organizations to arrive sooner, and other parties to become involved in disaster relief efforts.

In recent years, the Federal Emergency Management Agency (FEMA) has struggled to devote its limited resources to disasters across the United States of America [8]. By incorporating a software tool to identify disasters much more quickly than current conventional means, organizations such as FEMA could improve its own disaster relief efforts based on the severity of the discovered disaster. The effect of implementing this software could mean faster response times of authorities and rescuers, limiting disaster damage/cost, reducing the number of human casualties, and better allocation of finite resources. The ability to contribute towards improving disaster outcomes is motivation enough for this research.

The software requires the implementation of multiple tools to achieve the goal. The most critical component of this software is identifying disasters from tweets written by human beings. Natural language processing (NLP), is a subset of artificial intelligence that has the ability to process language and dialogue to analyze the nature of human communication. Natural language processing will be applied on the data that is provided by the Kaggle competition containing textual content of various tweets from Twitter users. These tweets, containing both disaster and non-disaster related content, can be used to train a model for identifying disasters in the real world. The performance of our natural language processing implementation will be gauged from the results of the Kaggle competition from which the data was provided. In addition, we hope use language processing to be able to expand our scope to classifying types possibly quantify the emotional impact of disaster from semantic value in tweets.

1. Related Work

We were already aware that multiple groups had attempted to use natural language processing to identify disasters through social media. The most notable example we found of existing research into the subject comes from [6]. Seeing as there is an active Kaggle competition on the same subject, we’ve decided to reference existing research and attempt to replicate or improve upon existing methods with the data from Kaggle because the data was hand-classified, providing a new dataset from which to gauge the effectiveness of the methods we intend to implement.

1. Background

The company Figure Eight wants to inspire future generations of data scientists to hone their skills for the benefit of everyone. They hand-classified a large but manageable-sized dataset of tweets as being related to disasters or not being related to disasters. This dataset was posted onto Kaggle with a soft ending on March 23rd, but will remain ongoing competition afterwards. The goal of the competition is to challenge people to find the best ways to quickly locate disasters through analysis of social media.

1. Methodology

Our methodology is very open-ended at the moment. We plan to utilize tools and methods described in some of our references such as [1], [5], and [6] based on ease of implementation and effectiveness. We might even be able to combine tools and methods to come up with a composite method for greater accuracy.

1. Timeline

We are currently at the end of week 4 of the semester. Our project reports are due at the end of week 12. During week 9 we hope to be testing our implementation of natural language processing because the start of week 10 is the end date for registering for the Kaggle competition. It will likely take until week 9 to refine our AI algorithm. During week 10 and 11, when Kaggle is less of a concern, we will refine the disaster categorization of our disaster detection algorithm, which is beyond the scope of the competition, and also create the PowerPoint presentation of our work thus far. Afterwards, we will prepare for a presentation at the Celebration of Scholarship.

REFERENCES

[1] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*, 3rd ed. Upper Saddle River: Pearson, 2016.

[2] M. Garcia, “How to Make a Twitter Bot in Python With Tweepy,” *Real Python*, 22-Sep-2019. [Online]. Available: <https://realpython.com/twitter-bot-python-tweepy/>. [Accessed: 06-Feb-2020].

[3] “Real or Not? NLP with Disaster Tweets,” Kaggle. [Online]. Available: https://www.kaggle.com/c/nlp-getting-started. [Accessed: 06-Feb-2020].

[4] F. Bravom, “felipebravom/AffectiveTweets,” *GitHub*, 19-Jun-2019. [Online]. Available: <https://github.com/felipebravom/AffectiveTweets>. [Accessed: 06-Feb-2020].

[5] D. Barker, “12 open source tools for natural language processing,” *Opensource.com*, 22-Mar-2019. [Online]. Available: https://opensource.com/article/19/3/natural-language- processing-tools. [Accessed: 07-Feb-2020].

[6] K. Stowe, M. Paul, M. Palmer, L. Palen, and K. Anderson, “Identifying and Categorizing Disaster-Related Tweets.” Boulder.

[7] S. Aslam, “Twitter by the Numbers: Stats, Demographics & Fun Facts,” Omnicore, Jan. 13, 2020. [Online]. Available: <https://www.omnicoreagency.com/twitter-statistics/>. [Accessed: 07-Feb-2020].

[8] E. Atkin, “America’s Natural Disaster Response Is Its Own Disaster”, The New Republic, Sep. 26, 2017. [Online]. Available: [https://newrepublic.com/article/145019/](https://newrepublic.com/article/145019/americas-natural-disaster-response-disaster). [Accessed: 07-Feb-2020].

[9]

[10]