

Classifying and Measuring Hate Speech in Twitter Using Topic Classifier of Sentiment Analysis



F. H. A. Shibly, Uzzal Sharma, and H. M. M. Naleer

Abstract The aim and objective of this research are to create a model to measure the hate speech and to measure the contents of hate speech. The descriptive analysis method of data science was used to describe and summarize raw data from a dataset. We used Twitter as the social networking Web site for this research to analyze and measure the hate speech and its classifications. A dataset from kaggle datasets was applied for this research. To produce statistical results, we used monkey learn machine learning libraries which are incorporated with Python program to design and develop a model to classify and measure hate speech and its types that could be trained and tested using sentiment analysis. Researchers have found that the majority of the tweets are based on racist and ethnicity, sex and religion-based hate speech are also widely available.

Keywords Machine learning · Hate speech · Social media and Twitter

1 Introduction

The rapid growth of social networking sites or social media such as Facebook, Instagram, Twitter, and other online social forums has transformed the pattern of communication and content publishing but is also increasingly misused for the circulation of hate speech and the organization of hate-based activities [1]. The term “hate speech” was formally defined as “any communication that disparages a person or a group based on some characteristics (to be referred to as types of hate or hate classes)

F. H. A. Shibly (✉) · U. Sharma
Assam Don Bosco University, Guwahati, India
e-mail: shiblyfh@seu.ac.lk

U. Sharma
e-mail: druzzalsharma@gmail.com

F. H. A. Shibly · H. M. M. Naleer
South Eastern University of Sri Lanka, Oluvil, Sri Lanka
e-mail: hmmnaleer@gmail.com

© Springer Nature Singapore Pte Ltd. 2021

D. Gupta et al. (eds.), *International Conference on Innovative Computing and Communications*, Advances in Intelligent Systems and Computing 1165, https://doi.org/10.1007/978-981-15-5113-0_54