Prevalent Disease Tweet Classification using Sentimental Analysis

Kevin Kurien Joseph[1], Rahul S[2], Siddhaarthan V[3], Indumathy[4]

Department of Computer Science and Engineering[1,2,3,4]

SRM Institute of Science and Technology, Vadapalani, Chennai, India

Abstract: *Interest in people's opinions about tweets has lately grown as a result of opinionated data that can be obtained on blogs and social media platforms. With everyone having access to the internet and the ability to tweet whatever they want, Twitter is one of the most widely used information sources in the world today. Therefore, there is a higher likelihood that people would be misled. It would be very helpful to come to a consensus on what the general attitude of the populace is, particularly during times of panic like the pandemic, in order to better grasp how ready the populace is to face such a potentially disastrous crisis. To handle these kinds of challenging problems, machine learning methods are typically applied. Analysing this kind of data manually requires more time. It is challenging to categorise opinions according to their polarity. This project was created specifically to analyse the moods of Twitter users using ML techniques and NLP tools to pinpoint the cause of COVID-19's reappearance. The perspectives of different people can diverge. After the system can classify the text's emotions, we train it using historical data. Machine learning methods are crucial to this project's ability to classify data. The Random Forest Classifier has also been chosen as the framework's preferred ML approach, with an accuracy rate of 88%. We assessed the framework using a variety of metrics, including F1-Score, precision, accuracy, and recall*

Keywords: precision, accuracy