INFORMATICS INSTITUTE OF TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

"SAIC"

Machine Learning Based Sentiment Analysis for Instagram Comments and Emojis

A Final Year Project by

Vivek Subendran

2018163 / w1715110

Supervised by

Mrs. Janani Harischandra

Submitted in partial fulfilment of the requirements for the BEng in Software Engineering degree at the University of Westminster.

05th May 2022

ABSTRACT

More and more people use emoticons in their texts to communicate their feelings or remember their comments. Previous machine learning algorithms mainly focused on categorizing text, emoticons, or images and ignored emoticons with text, resulting in many emotions being skipped. The study provided a sentiment analysis algorithm and approach that used text and emoticons. Both data forms were evaluated in this study using machine learning to detect sentiments from Instagram comment data using numerous features such as TFIDF, N-gram, and emoticon lexicons. This study shows that when using emoticons, the associated sentiment outperforms the sentiment represented by the analysis of text data.

Keywords - Machine Learning, Naïve Bayse, Data training, Emoticon, Sentiment Analysis

Subject Descriptors – Computing Methodologies > Machine Learning > Natural Language Processing