



INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : Loitongbam Gyanendro Singh

Roll Number : 176101012

Programme of Study : Ph.D.

Thesis Title: **Sentiment analysis of tweets on societal topics.**

Name of Thesis Supervisor(s) : Dr. Sanasam Ranbir Singh

Thesis Submitted to the Department/ Center : CSE

Date of completion of Thesis Viva-Voce Exam : 11/11/2021

Key words for description of Thesis Work : Sentiment Analysis, Deep learning, Multi-view learning

SHORT ABSTRACT

With the increasing popularity of Twitter, a surge in understanding public opinions on various social issues has been evident. One of the parameters often considered in such studies is understanding public sentiment toward target policies or issues. Earlier studies on sentiment analysis primarily focus on commercial domains such as product review, movie review, restaurant review, etc. However, sentiment analysis of opinions on the societal domain faces unique challenges because of a wide range of topics with diverse vocabularies, target aspects, and different language constructs. This thesis aims to investigate the characteristics of tweets in societal and non-societal domains and build an effective sentiment analysis classifier for the Societal domain. Unlike regular texts, sentiment analysis of tweets needs to deal with various challenges. Tweets are generally short and often under-specified due to character limits. They often consist of noises due to informal writing (shorten/elongated text), misspelling, multilingual code-switch, and code-mixed contents. Contributions are made to address the challenges of sentiment analysis of tweets such as under-specificity, noise, and multilingual content, by proposing deep learning-based methods to improve the performance of the sentiment classifiers in the societal domain. We proposed a method to represent tweets in a heterogeneous multi-layer network by incorporating the relation of keywords, hashtags, and users to enrich the representation of tweets by adding information relevant to the tweets, such as keywords, hashtags, and users information. The proposed method for representing tweets in a network structure can capture better information than text-based representation by adding relevant related information. Further, we proposed a multi-view learning framework to incorporate both text and graph views of a tweet to enrich the tweet representation for the sentiment classification task. It is observed that the proposed multi-view learning framework can better perform classify tweet's sentiment than its single-view counterparts.