



**INFORMATICS INSTITUTE OF TECHNOLOGY**

**In Collaboration with**

**UNIVERSITY OF WESTMINSTER (UOW)**

BEng (Hons) in

**SOFTWARE ENGINEERING**

**SORTING AND GROUPING TEXT MESSAGES IN  
SOCIAL MEDIA GROUP COMMUNICATION**

for

Final year Project 2017/2018

By

**Durantha Kalpage**

2014234 | w1560224

Supervised By

**Mr. Guhanathan Poravi**

©The copyright for this project and all its associated products resides with Informatics Institute of  
Technology

## Abstract

With the development of the social media communication, it's become more popular among teenagers to exchange their ideas and thoughts. As per the development Group communication is, using many to many group communication, there is a tense to lose important messages or notice due to flood of messages. It's not easy to find out a message from a particular person to a group by searching the entire chat history. In this research, author reviews about the classification technique and methods to sort incoming messages in order to overcome the message overloading.

The target audience of the project is the teenagers who are using social media communication along with their day today activities. All the incoming group messages will be sorted automatically according to the extracted keywords from the message body. Users will be able to get messages sorted automatically in to sub topic/categories. Users have the ability to view selected topics which is more efficient than reading an entire group chat history.

All task of this project were built on Flask web service and text mining is processed using Natural language tool kits to extract and manipulate the text bodies of the incoming messages. Automatic topic generation were done using the machine learning principles and clustering algorithms. PhpMyAdmin is used as the database and Telegram message platform is used to get data set for the relevant activities.

**Keywords:** Message Overloading, Text mining, data sets, Machine learning, Clustering