Characterizing and Countering Communal Microblogs During Disaster Events

In this paper author is describing concept to detect communal hate tweets spread in social media networks during disaster events occurred. Sometime during natural disaster event, peoples may use social media networks to spread current situation or relief activities happening at disaster area and some corrupt peoples may use this situation to spread hate messages to disturb peace.

To detect such hate tweets author is using rule base concept to detect such tweets, in this rule we will find out that such hate tweets may get more re-tweet counts and may have some hate words such as Muslims, Christians, terror, attacks etc and we look such words from tweets to define or classify as communal tweets and tweets not contains such words may be consider as non communal.

Author is saying we can download such hate words for religion and race from different web sites such as <http://www.translationdirectory.com/glossaries>, www.hatebase.org etc.

To implement above concept I downloaded tweets dataset from <http://www.cnergres.iitkgp.ac.in/disasterCommunal/dataset.html> wbsite

First we will upload tweets dataset and then build train classifier by assign each tweets to either communal or non communal class by applying rule concepts and by using hate words.

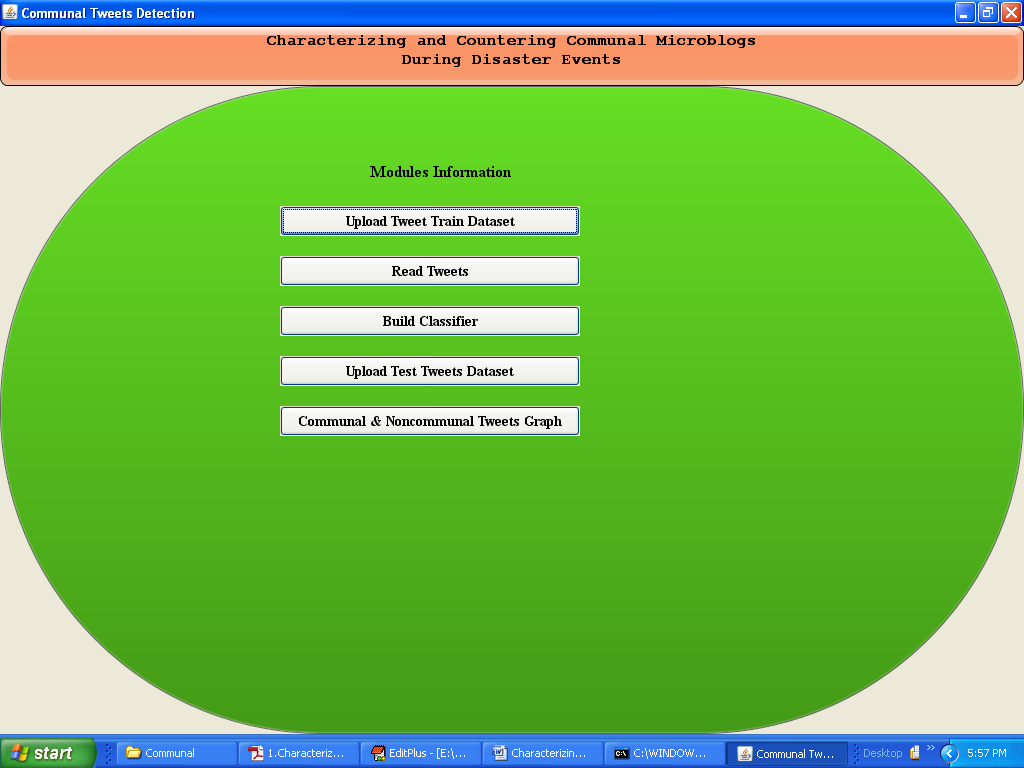
Rule concept is to check whether tweets contain communal hate words or not. Test tweets will be apply on train classifier to predict whether uploaded tweet is communal or not communal. Some tweets may contain hate words but it will consider as anti communal tweets which try to stop communal hate tweets. Example of anti communal tweet is

‘Threat from a kashmiri Muslim not a terrorist. Every1 should keep this as proof’

In above tweet we can see it contains terrorist word but it’s not a communal tweet but it’s an anti communal tweet.

Features.txt contains all hate words list for communal tweets

Screen shots



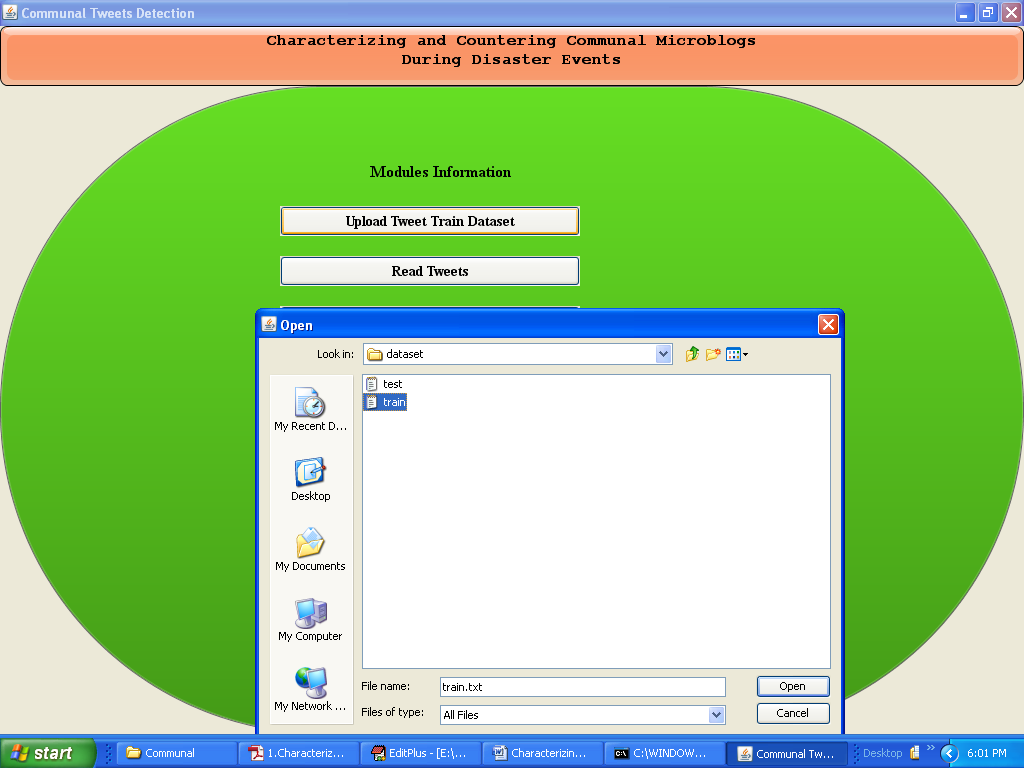
Click on ‘Upload Tweet Train Dataset’ button to upload train tweets and this train tweets will not have communal or non communal class, application will identify by using rules and hate word list. Tweets example from train dataset

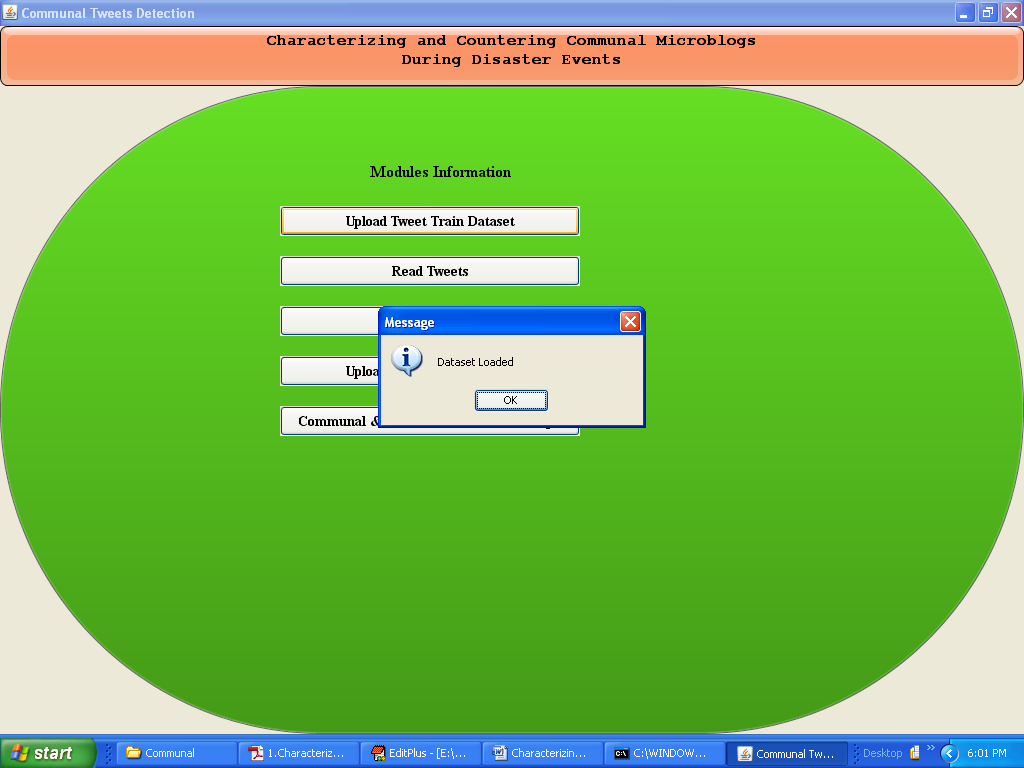
Retweet, username – tweets text

56,gkrajvanshi - Huh, its muslim behind California attack

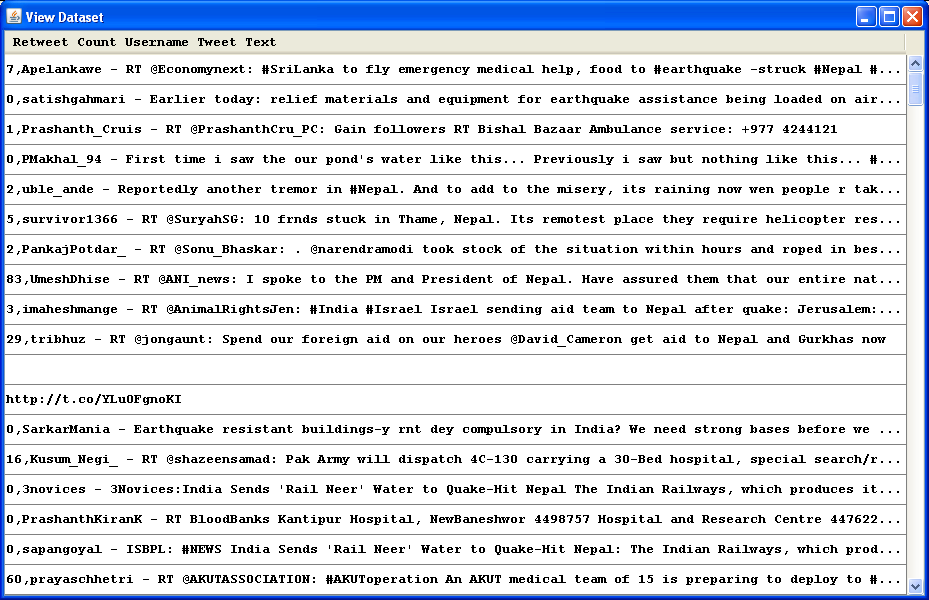
56,gkrajvanshi - Threat from a kashmiri muslim not a terrorist. Every1 should keep this as proof

In above tweets 56 is the re-tweet count and then username and then tweet text will be there but there is no class definition

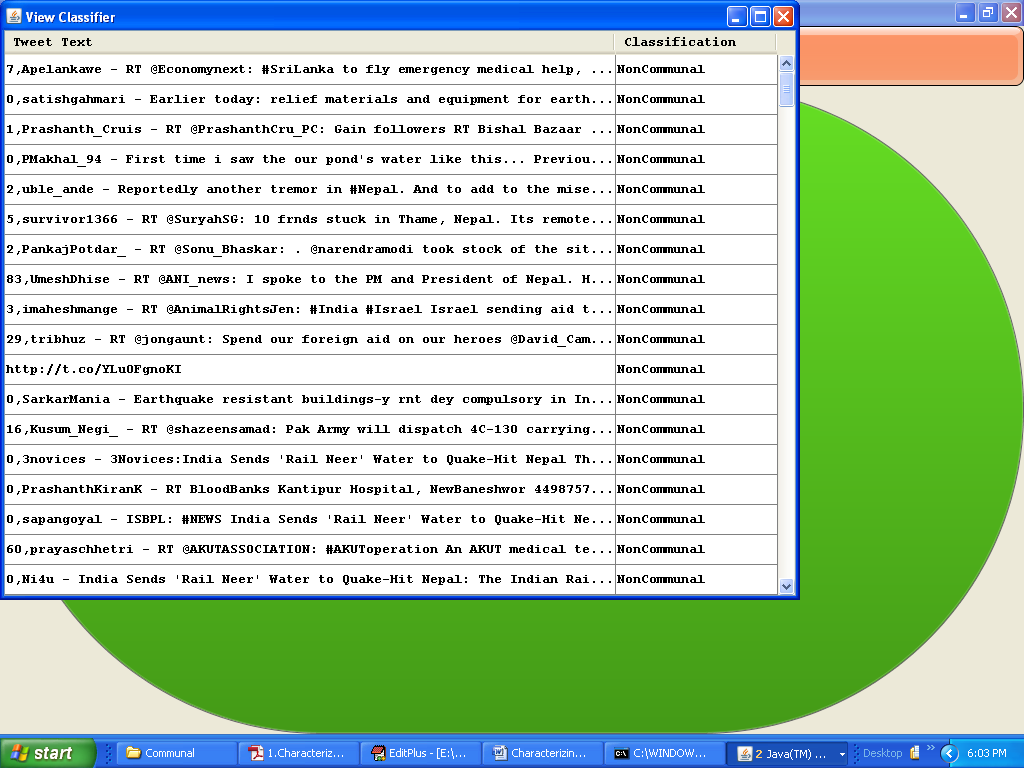


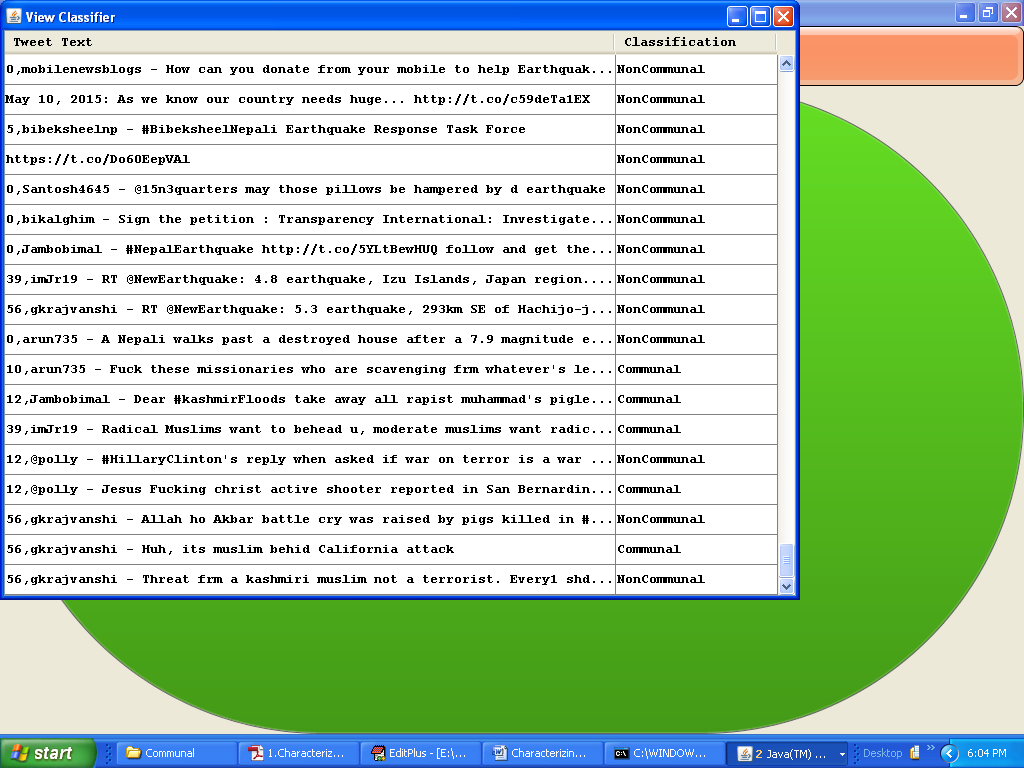


Now click on ‘Read Tweets’ button to read all tweets



In above screen we can see 7 is re-tweet count and then username and then tweet text. Now click on ‘Build Classifier’ to assign class to each tweet



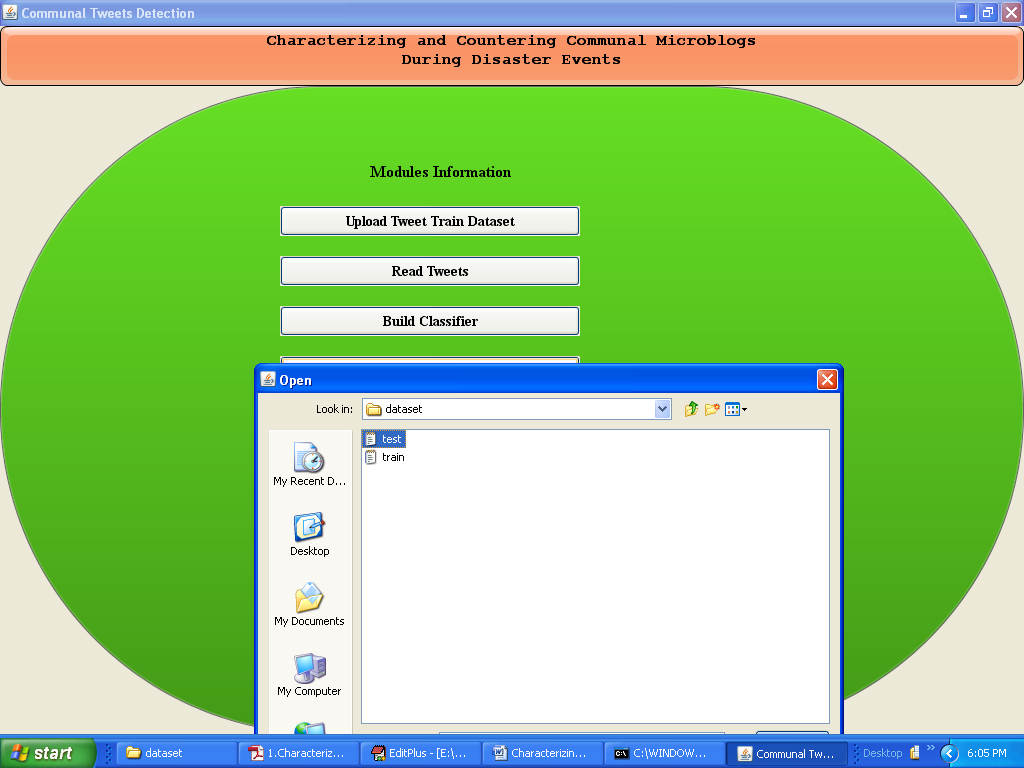


In above screen we can see each tweet assign with class. Now upload test dataset and predict the class name from train dataset. Examples of test tweets are

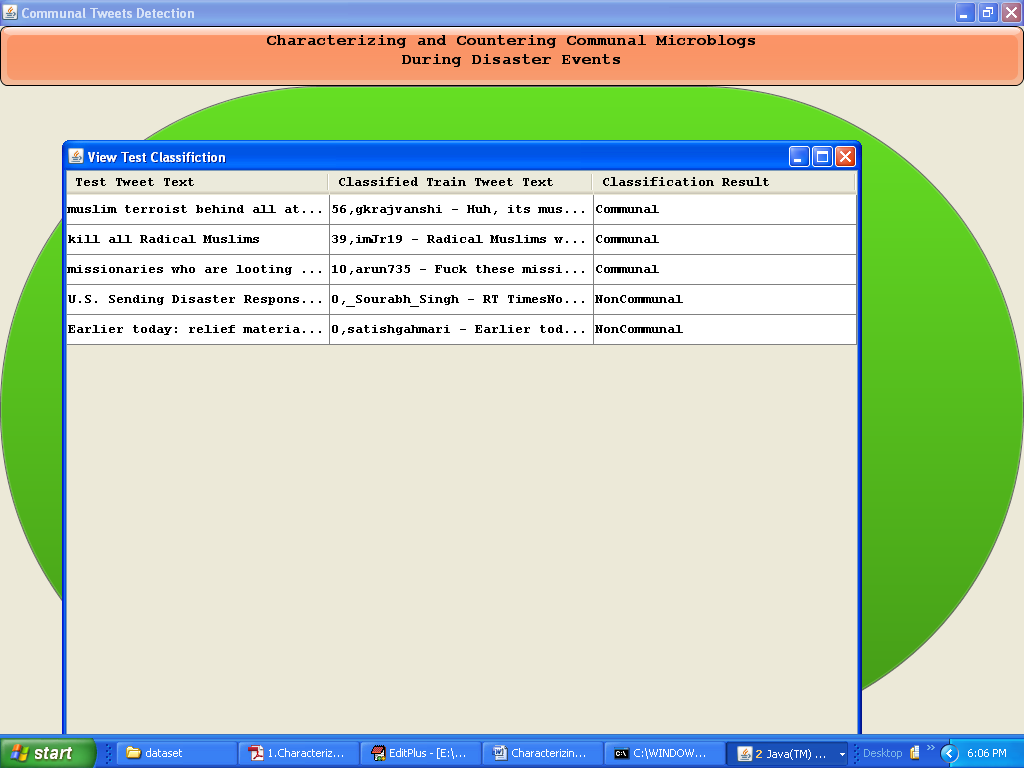
“muslim terroist behind all attack

kill all Radical Muslims

missionaries who are looting frm whatever's left after earthquake”

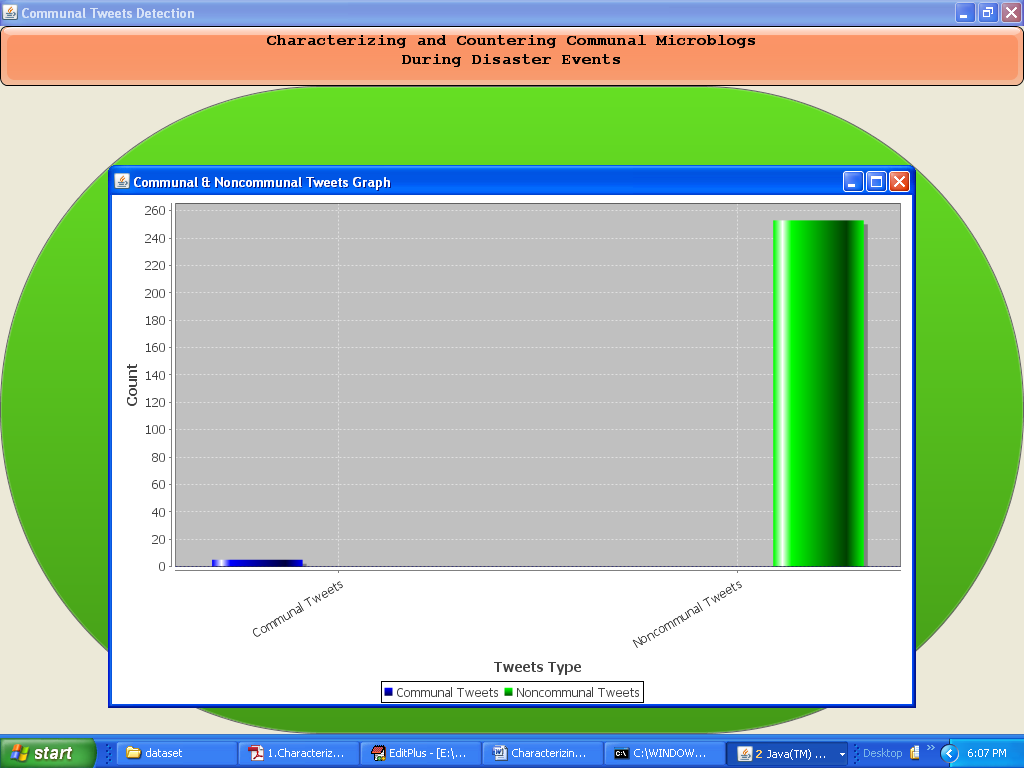


After uploading test dataset will get below screen



In above screen we can see test tweet and classified train tweet with predicted value as communal or non communal.

Now click on ‘Communal & Non communal Tweets Graph’ to see total number of communal and non communal tweets



In above graph x-axis represents tweet type as communal or non communal and y-axis represents count