**Abstract**   
Microblogging today has become a very popular tool for communication and voicing opinions among Internet users, where millions of users share opinions on varied topics. Microblogging platforms like Twitter allow users to share views within a 140 character limit, leading to a high rate of information compression. Users also have the option of adding images, links and videos among other things, which makes the type and variety of content very diverse. Here processing the tweet involves extraction of metadata of tweet, geocoding the physical address in a tweet, analysing the sentiment of content in the tweet text and extracting the significant and key phrases from a text. This project mainly focuses on performing three tasks. First is to collect the tweets from Twitter, having a chosen keyword, save them on a database and clean them up to have only necessary information. This is achieved using the Twitter Application Program Interface (API) along with Python language. Secondly, classify the tweets based on Sentiment Analysis, which determines the judgment or evaluation of a user with respect to the chosen topic. This is performed using Python language. Third is to represent the locations and frequencies of the tweets geographically. The details of the users of the corresponding tweets are collected using REST and Streaming APIs provided by Twitter and the geographic representation is acheived using the packages of R language.

**Contents**

**Chapter 1   
Introduction**The growth of the computer age has brought along with it a very large amount of information that is available for use . This has gained a lot of traction especially in the past few years, with scientists and the general public alike looking for better ways to utilize and make sense of the existing data. The existing patterns of communicating data, which mostly include listing out the data might not always be the best method to ensure that the information behind the data is conveyed efficiently. One of the better methods used in this regard is the process of visualization.

Visualization is gone for human comprehension in preparing the data proficiently and successfully. The quickened development of 'informal communities' (Example,Twitter) makes conceivable, to exchange and share data to numerous client's quick with less cost. The potential result of interpersonal interaction encourages a client to reach and cooperate a great many different clients. Organizations are building Third gathering applications, which are exploratory in conveying apparatuses to advantage client. It examines the conclusions, client sees, new thoughts, open hobbies, and their engaged exercises of a huge number of client round the globe. Showcasing firms additionally get included in dissecting client inputs and practicing over open supposition, and the brake out of most recent patterns in the masses in overhauling the items and administrations. The crude material in building the outsider application is mass volumes of information that needs to procedure to get data. The extraction of data from crude information put additional weight on applications that impedes viable usage of accessible information. Content examination might likewise allude as content digging for content investigation, to enhance quality perseverance and adds sense to the importance of information. The work proposed visualizes a structure that not only reduces the amount of existing noise that usually comes along with the corpus, but also looks at methods that can ensure that the amount of noise be reduced to a great extent.   
  
This paper expects to assess the accessible API's to get access information from Twitter, and execution of suitable methodology to fabricate database of interpersonal organization information (twitter). To make it helpful for perception of twitter information, it is imperative to comprehend which is proficient and compelling in usage and support. Additionally inspected and thought about existing gazetteers and Entity extraction libraries. For an undertaking of actualizing NER (Names Entity Recognition) to concentrate annotation particular to the characterized examples and organizations after legitimate investigation of information. Assumption examination have knowledge to distinguish the positive and negative sense in the content, the assessment concentrates for the most part on the conduct viewpoints and words or expressions that implies the human feelings.

* 1. **Literature review**
  2. **Problem definition**
  3. **Organization of project report**

**Chap 2 : designing part**

**9 uml diagrams**

**System model**

**Conclusion**

**Basepaper ( reference )**

**Bibliography**