Sentiment Analysis on Social Media

Group No. 14

Riddhi Prajapati(Leader)- 15104015

Pratiksha Patil - 16204011

Aafreen Shaikh - 16204009

Project Guide- Prof. Sunil Sushir

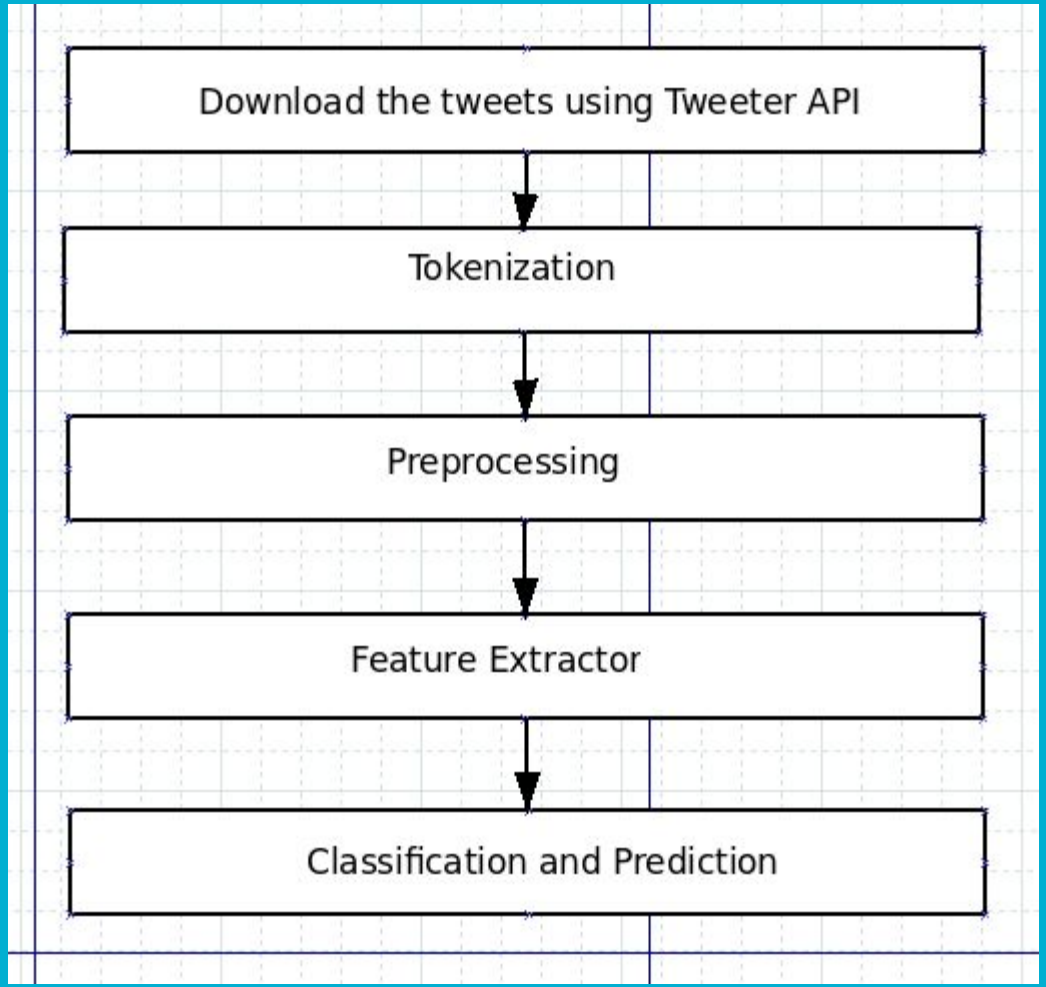
Contents

Abstract

The time spent by uses is almost two or more hours looking for papers that asks search engine to optimize the result.Sentiment analysis is a method which analyses emotions or views of an individual regarding to a topic. Social media in today's world is a collection of people's feedback, emotions, reviews and even personal experiences which result in proper knowledge of the given content. Here we design a framework for sentiment analysis in case of twitter data. The available data is analyzed in terms of positive or negative and an clear summary is obtained. The input data that needs to be analyzed is given as input to the NLP and processed.

Introduction

- Sentiment analysis is the process of determining whether a piece of text is positive, negative or neutral.
- The text may be a sentence, a tweet, an SMS message, a customer review ,a document and so on.
- R programming language will be used for data analysis along with NLP algorithm and Lexicon approach.
- We are performing Sentiment analysis on twitter data, because it is:
    1. popular microblogging site .
    2. 240 plus million active users.
    3. 500 million tweets are generated everyday.
    4. Twitter audience varies from common man to celebrities.
    5. Users often discuss current affairs and share personal views on various subjects.


Download the tweets using Tweeter API
Tokenization
Preprocessing
Feature Extractor
Classification and Prediction

Literature Review

Sr. No	Author Name	Title Of Paper	Tools/Algorithm	Issues	Overcome
1	Ms.K.S.Taranya	Onto-based Sentiment classification using Machine Learning Techniques.	NLP, Statistics, R-programming Language	In preprocessing stage, various rules and regulations are defined to standardize the text that make text mining process efficient.	The machine learning algorithms are applied for better results.
2	Md.Kashif Hanif, Fakeeha Fatima	Text Mining: Techniques, Applications and Issues	Classification, clustering, k-means clustering, NLP.	Domain knowledge integration, varying concepts granularity, multilingual text refinement, NLP ambiguity.	To develop specific algorithms to overcome the issues.

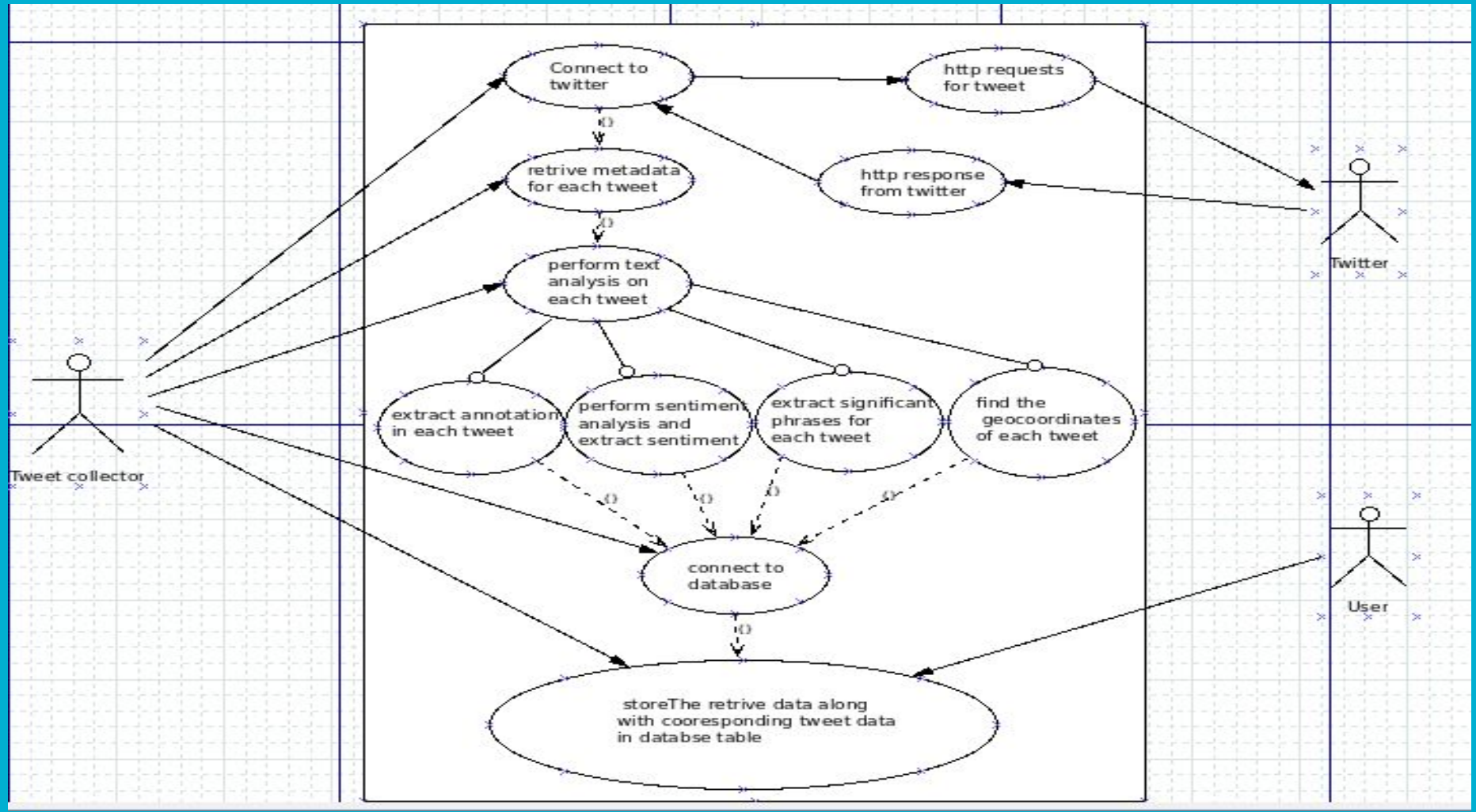
Sr. No	Author Name	Title Of Paper	Tools/Algorithm	Issues	Overcome
3	Kudakwashe Zvarevashe	A Framework for Sentiment Analysis with Opinion Mining of Hotel Reviews	Opinion mining, Sentiment Analysis, Machine Learning algorithm , Natural Language Processing(NLP)	Issues arise that some comments may be wrongly viewed as neutral	Here lot of research is done in fine tuning the feature extraction algorithm.
4	Andreea Salinca	Business reviews classification using sentiment Analysis	Naïve Bayes for text classification. TF-IDF algorithm.	The accuracy of the system decreased with 14% when compared to first approach result. Here also Sentiment ambiguity occurs.	The performance can be improved by the usage of bigrams or trigrams, word chunks, or part-of-speech as features in order to distinguish the same word features that use different POS.

Problem Statement

A major benefit of social media is that we can see the good and bad things people say about a particular brand or personality.The bigger your company gets difficult it becomes to keep a handle on how everyone feels about your brand.For large companies with thousands of daily mentions on social media, new sites and blogs, it is extremely difficult to do this manually.

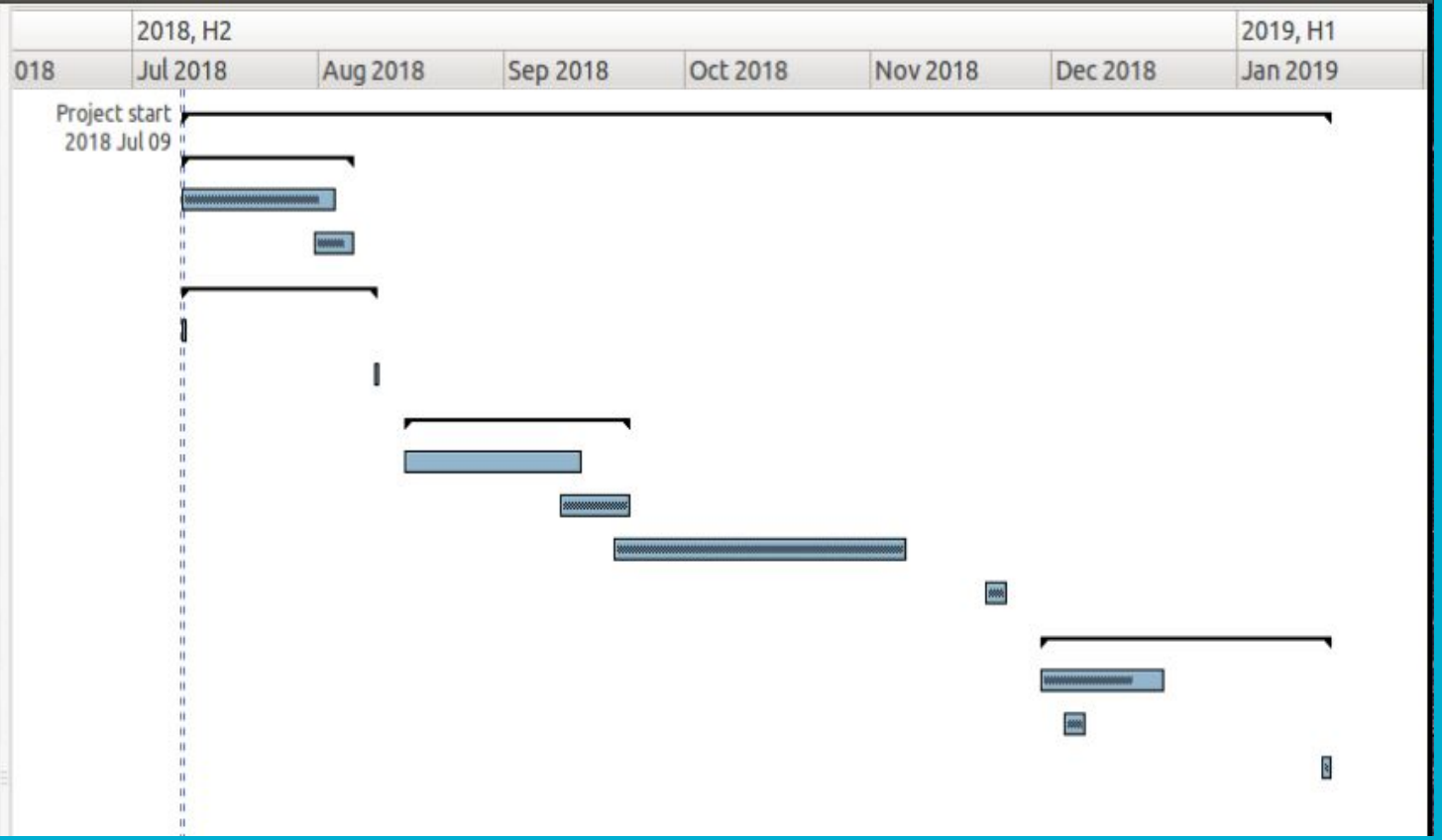
Project Scope

- This project will be helpful to common people, companies.
- It will be helpful as we all can get reviews about a newly released product by analyzing the comments or reviews posted by people.
- Also the particular content can be analyzed whether it is depicting a positive, negative or nuetral report.
- Hence by analyzing the tweets analyzer can get result on how positive or negative or nuetral people are about it.


Connect to twitter
http requests for tweet
retrive metadata for each tweet
http response from twitter
Twitter
perform text analysis on each tweet
Tweet collector
extract annotation in each tweet
perform sentiment analysis and extract sentiment
extract significant phrases for each tweet
find the geocoordinates of each tweet
connect to database
User
storeThe retrive data along with cooresponding tweet data in databse table

Project Future plan

WBS	Name	Work
1	**Sentiment A**	**119d**
1.1	**Requireme**	**25d**
1.1.1	System R	20d
1.1.2	Customer	5d
1.2	**Planning**	**2d**
1.2.1	Time	1d
1.2.2	Budget	1d
1.3	**Design**	**32d**
1.3.1	Design AF	22d
1.3.2	Design GU	10d
1.4	Coding	35d
1.5	Testing	4d
1.6	**Presentati**	**21d**
1.6.1	Feedback	15d
1.6.2	Proposec	4d
1.6.3	Final Pres	2d



Summary

- People have always had an interest in what people think, or what their opinion is. With the interesting numbers of people are using websites and services to express their opinion.
- With social media, it is becoming easier to automate what public opinion is on a given topic, news story, product, or brand.
- Opinion that are analyzed from such services can be valuable.datasets that are gathered can be analyzed and presented in such a way that it becomes easy to identify if the online mood is positive,negative or neutral.
- This allows individuals or business to be alert when a negative opinion arises, alternatively, positive sentiment can be identified thereby allowing the identification of product advocates or to see which parts of a business strategy are working.

References

- Onto-based sentiment classification using Machine Learning Techniques.Walaa Medhat, Ahmed Hassan, HodaKorashy, “Sentiment analysis algorithms and applications:A survey”,2014.
- R.Ezhilarasi, R.I.Minu, “Automatic Emotion Recognition and Classification”, International Conference on modelling,optimization and computing,2012.
- B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," Proceedings of the ACL-02 conference on Empirical methods in natural language processing.
- S. M. Weiss, N. Indurkhya, T. Zhang, and F. Damerau, Text mining:predictive methods for analyzing unstructured information. Springer Science and Business Media, 2010.

Thank You !!!