International Journal of Computer Engineering in Research Trends
Multidisciplinary, Open Access, Peer-Reviewed and fully refereed
Research Paper Volume-6, Issue-11, November 2019 Regular E-ISSN: 2349-7084

Sentiment Analysis on Social media

A. Manjula[1*], Dr. A. Rama Mohan Reddy[2]

[1*] *Computer Science and Engineering, Sri Venkatewsara University College of Engineering, Sri Venkateswara University, Tirupati, India*
[2] *Computer Science and Engineering, Sri Venkatewsara University College of Engineering, Sri Venkateswara University, Tirupati, India*
e-mail: manjulaanumula777@gmail.com, ramamohansvu@yahoo.com
**Corresponding Author: manjulaanumula777@gmail.com,*

Available online at: http://www.ijcert.org

Received: 30/11/2019, Revised: 05/12/2019, Accepted: 17/12/2019, Published: 27/12/2019

Abstract: - This paper suggested different improvement process to track and prevent shoplifting using technologies – Artificial Intelligence, Machine Learning, Data Analytics, and Process Automation. The detailed work is done in developing a system used for opinion analysis of a product or a service. The system readily processes the tweets by pulling data from twitter posts, pre-processing it and connecting it to Twitter API by REST call method and showing it graphically. We have given the analysis for the public tweets by API and filter them for various products, persons and services. For written product reviews, the best solution is a video review. Collecting comments from YouTube videos and extracting the exact tone or behaviour behind it. The most widely used approaches in opinion mining focus only on tweets or written product reviews available on websites like Amazon. Various emotions that can deal here namely Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Curiosity, Excitement, Gratitude, Serenity, Hope, Pride, Amusement, Jealousy, Guilt, Discouragement, Frustration, Rejection, Disappointment, Loneliness, Interest, lack of interest, Concern, Sympathy and Calm. Online news is also now trending and extracting the proper tone behind the news. By applying machine learning algorithms like feature extraction, classification algorithm, natural language processing techniques, and more, polarity can be identified. The analysis is used to classify the sentiment as positive, negative, neutral, strongly positive, weakly positive, strong negative, and weak negative. The results have shown textually and graphically.

Keywords: Sentiment analysis, Twitter, YouTube, Online-News, data, analysis.

---

1. Introduction

As technology increases day by day, vast volumes of data also increase. Over a decade ago, a GigaByte data generated per day, but from now on it were in seconds. Time changes everything; the world is moving through developing side more and more. Different departments generate data vastly, such as Institutions, Companies, Social media sites, Hospitals, Companies, Governments, etc., among these 99 out of 100 uses Social media sites all over the world. Various social media sites like Twitter, YouTube, Facebook, Whatsapp, Instagram, Snap Chat, etc., Sentiment analysis is also known as Opinion mining. It is the method of finding out the tone behind words. It helps to gain an understanding of the attitudes, opinions and emotions expressed within an online mention. Opinions, feedback and critiques provided by internet users show attitudes and sentiments towards specific topics, products, or services. It is difficult to read and understand the massive volume of data. Thus, social media sentiment analysis plays an essential role in solving and making better decisions. Various social media sites that are dealing are Twitter, YouTube, and Online news. From multiple sites, collecting different data-sets and doing transformations and obtaining the results defining the emotional status. Sentiment analysis comes under Natural Language Processing. The sentiment is an attitude or an emotion or a feeling. It mainly studies about users sentiments towards certain products or persons or services.

With the help of the internet, people able to post their views, points and their content through social media sites such as forums, micro-blogs, online networking sites. The social media sites that are dealing here are Twitter, YouTube, and Online-news. Every website has its pros and cons — over 1.5 million tweets generated per day by users. Every famous person like political leaders, actors, actresses, company holders, Institutions, uses Twitter a lot to post their recent activities and views. To know the emotional status of tweets , Twitter data extracted and do some transformations and analyze them and categorize them into the positive, negative, neutral, actively positive, strong negative, weakly positive, weak image. Any twitter user can extract the public data from Twitter using Application Programming Interface (API). Twitter currently has mainly three different API's such as the REST API, the Search API and the Streaming API. Developers can able to gather status data and user information by using REST API. While the Search API allows developers to query specific Twitter content. Whereas, the Streaming API able to collect the real-time data on Twitter. To achieve better results, API's can be mixed.

Not only Textual data but also videos can be extracted and analysed. YouTube videos are famous video sites for reviews. Collecting video comments from different users and can be decomposed them into positive, negative and neutral. By this, it is easy to define that is video is worth it or not. Anything in this society is news. Before the 19th century, typescripts, newspapers which were used to know about the news. But now, technology increases, news can be read, knew through the Internet, which is also known as Online-News. Various news articles related to different regions can be notified, analysed and categorized them into positive, negative and neutral based on the polarity value.

2. Related Work

The primary source of generating real-time data by individuals, not only producing but also accessing and sharing data is Social media. Theirs exists various data sets to check polarity and subjectivity classification (sentiment analysis). Here the complexity lies in identifying the discrete emotion behind the data sets. Analyze, classify and investigate the relationship between emotions and reporting them [1].

Different users from different communities, countries and languages communicate with each other using a platform called "Social Networking Sites". These sites become a significant source for generating, collecting public information and can serve as an insight in taking better decision [2].

Online videos have become more prevalent in the daily world. The natural evolution of written reviews is video reviews. YouTube has become popular in reviews through commenting. Opinion analysis can be carried out in order to obtain fine-grained performance [3].

Online mental health forums help individuals in finding the progression of emotional

states, and that can be studied through the individual's posts. From that estimation can be done and vary person to person [4].

Over a decade, there is a full demand for sentiment analysis. There is growth in algorithms, data sets, models and environment. It is difficult to understand how a particular model fits over specific tasks and domains. Getting a better understanding of how different model architectures work on different data sets [5].

It is essential in getting a machine that understands the exact meaning of language. It varies from different fields. Its main goal is to train the model to label-comment sentiment along with polarity and subjectivity with emoticons [6].

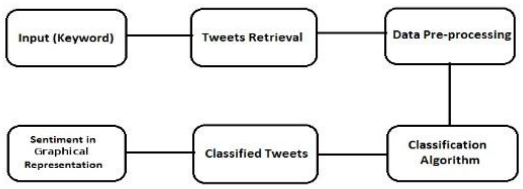
The Internet plays a significant role in our daily life. Everyone has a social media account not only to share their data but also to access and to generate data through emotions. Finding out the better decision from emotion intensities collected from Twitter and even predicting the real-time value for it.[7]

3. Methodology

For this work, we have been used following system specifications such as hardware requirements for the job is Intel core i5 or i7 processor, Minimum 3 GB RAM, 8 GB RAM is recommended, Minimum 500 GB of hard disk space similarly, following are the Software Requirements which are operating system either windows 8/ windows 10 (32 bit/ 64 bit) ,Programming language R,Python ,Software version - version 3.33(R),version 3.8(Python),Software's used- Rstudio,Sublime Text,Python IDLE.

Implementation:

The algorithm followed through these steps in various social media sites.



I. **Twitter:-**

Steps

- To create a Twitter Application
- Authenticating Twitter Application with login credentials
- Creating the function to search and returns the data through Twitter API
- Installing the packages and libraries

II. YouTube:-

Getting YouTube data and analyzing them can be carried out in the following steps:

- Authentication: - open Google API console in browser and create your API service using several API tools like YouTube Analytics, Data Analytic V3.
- Create one with named "My project".
- Credentials - API_Key= "Aiza*******"
- After successful completion of this procedure, YouTube videos respective id's has to be noted.
- Getting YouTube id's:- select a product and search in YouTube, it will display the corresponding related videos to it.
- Search about any topic, product or service (example: Narendra Modi, Samsung Galaxy, iPhone 11, K20 pro)
- Know the appropriate sentiment from the comments about the product.
- Processing and getting output:- get the ID of the video which to extract
- Perform various operations on the data collected and the output will be displayed in terms of plotting.

III. Online-news:-

Getting online- news and analyzing them can be carried out by the following steps:

- Open any python editor and execute the code as there is no authentication of API. It just requires the link of the URL of the online-news(here, the Hindu)
- It displays various Regional, National, International, Science and Business news.
- Regional news is from Karnataka, Tamilnadu, Andhra Pradesh, Kerala, Telangana.
- At once it carries 20 different news related to single entity and corresponding

transformations performed to find out the polarity(sentiment behind it)

- News, Sentiment value corresponding to a particular entity and Overall news sentiment can be displayed.

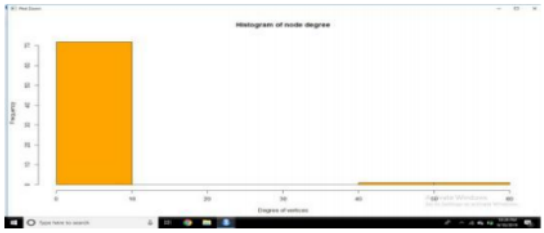
4. Results and Discussion

It includes screenshots of the output, graphs, plots and the description behind it.

Output for YouTube data

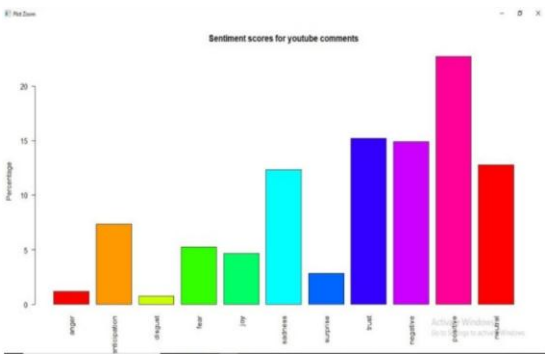
Network diagram for the nodes on that particular video

Output for YouTube data



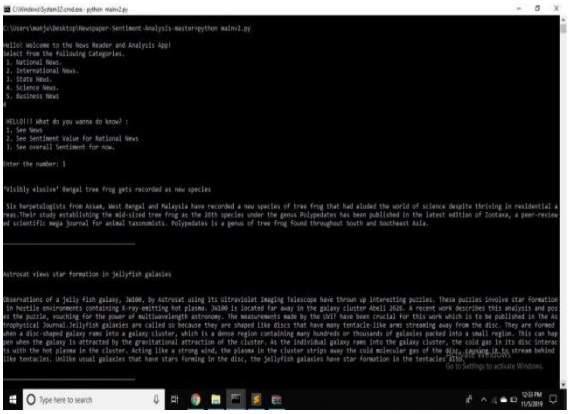
Histogram for displaying the nodes and its frequency over the network

Output for YouTube data



Sentiment scores for YouTube video comments

Output for Online-News data

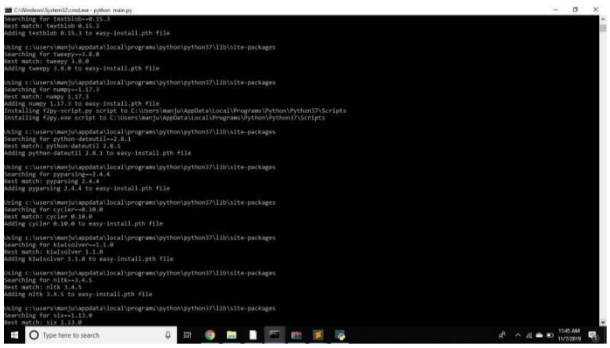


Displaying the various Regional, National and International news

Output for Twitter data

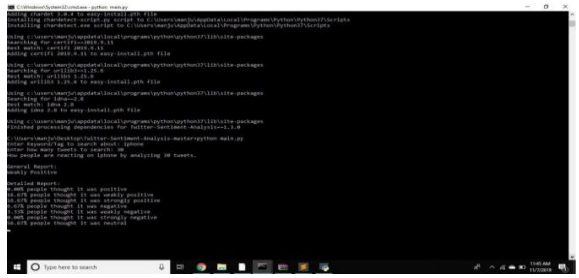
Installing packages related to twitter sentiment analysis

Output for Twitter data



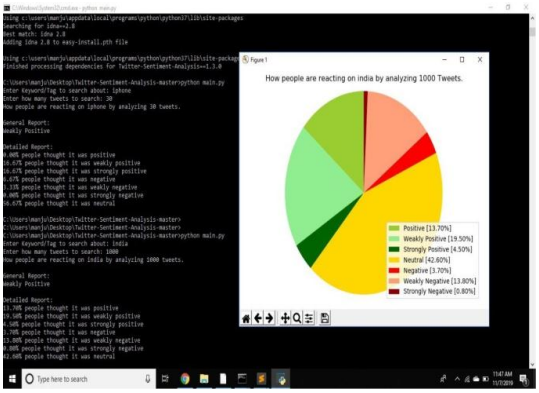
Importing packages for the execution of a program

Output for Twitter data



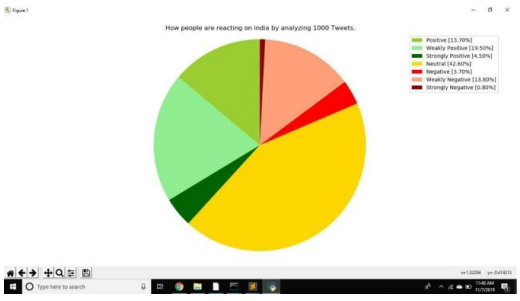
Output for Twitter sentiment analysis, displaying the detailed report along with percentages showing positive, negative, weakly positive, strongly positive, strongly negative, weakly negative and neutral for the keyword- "iPhone".

Output for Twitter data



I am showing the detailed report along with the pie chart (in percentages).

Output for Twitter data



The pie chart shows different sentiment scores in different colours. Each colour defines its sentiment and its percentage. This is a pie chart plotted for the search keyword-"India" and collected for "1000" tweets.

5. Conclusion and Future Scope

This paper has discussed the digital the proposed model overcomes the drawbacks of the traditional survey methods. The system gathers the necessary data from customers/people by using social media sites like Twitter, YouTube, and Online-News. These help to stream the public data alone. From this data, various pre-processing operations done and generate reports such as histogram, graphs, bar plots, pie-charts, etc., Real-time analysis is a powerful tool for business, marketing, politics, which has a tendency to extract the result of current trends through feedback and helps in decision making. In addition, this model can be faster with parallel distributive databases. Also, many high- end report representations can be used for much more in-depth analysis. Future work involves:

- Mining fine-grained opinions on closed captions of YouTube.
- Videos Deep emotion intensities.
- Fine-grained emotions analysis.

Sentiment analysis can be used in various fields and has many applications used are:- Product reviews, Movie reviews, Getting customer needs, accessible accessing in online news data, Spam detection and more.

References

[1] Hendrik Schuff, Jeremy Barnes, Julian Mohme, Sebastian Pado, and Roman Klinger," "Annotation, Modelling and Analysis of Fine-Grained Emotions on a Stance and Sentiment Detection Corpus " by", Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social media Analysis, pages 13-23, Copenhagen, Denmark, September 7-11, 2017.

[2] Sunidhi Sharma, D.K.Sharma, Supriti Sharma, "Text Analysis and Sentiment Analysis using Facebook in R Language: Case studies", International Journal of Computer and Mathematical Sciences, ISSN 2347-8527, vol 6 Issue 12, December 2017.

[3] Edison Marrese-Taylor, Jorge A. Balazs, Yutaka Matsuo, "Mining fine-grained opinions on closed captions of YouTube videos with an attention-RNN" Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social media

Analysis, pages 102-111, Copenhagen, Denmark, September 7-11, 2017.

[4] Kishaloy Halder, Lahari Poddar, Min-Yen Kan, "Modeling Temporal Progression of Emotional Status in Mental Health Forum: a Recurrent Neural Net Approach" Proceedings of Workshop on Computational Approaches to Subjectivity, Sentiment and Social media Analysis, pages 127-135, Copenhagen, Denmark, September 7-11, 2017.

[5] Jeremy Barnes, Roman Klinger, and Sabine Schulte im Walde, "Assessing State-of-the-Art Sentiment Models on State-of-the-Art Sentiment Datasets" Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social media Analysis, pages 2-12, Copenhagen, Denmark, September 7-11, 2017.

[6] Keenen Cates, Pengcheng Xiao, *, Zeyhang, Calvin Dailey, "Can Emoticons Be Used To Predict Sentiment?" Journal of Data Science 355-376, April 04, 2018.

[7] Prabaharan Poornachandran, "deepCybErNet at EmoInt-2017: Deep Emotion Intensities in Tweets" Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social media Analysis, pages 102-111, Copenhagen, Denmark, September 7-11, 2017.

