International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 10 | Oct 2021 www.irjet.net p-ISSN: 2395-0072

Sentiment Analysis of Social Media Comments using Data Mining

Diksha Jasuja[1], Mrs. Aman Paul[2]

[1]M.tech, CSE, Jannayak Chaudhary Devilal Memorial College of Engineering
[2]Asst professor (CSE), Jannayak Chaudhary Devilal Memorial College of Engineering

---

Abstract - The *worldwide presence of the internet attracts the wide range of users. Because of it a range of social Media network has come into the existence. There are various social media tools like blogs, twitter, face book which are very popular and also very effective. A massive amount of data generated by their users is difficult to handle all the data manually. Thus to make this task easy an automatic approach is used to analyze and categories the data for the social network. For this data mining technique is used*

Key Words: Social media, Data Mining, Social Media Analysis, Data Mining Techniques, Social Network.

1. INTRODUCTION

1.1 Social Media

Social media acquires online communications channels dedicated to community-based input, interaction, content-sharing, and collaboration. Facebook, Twitter, Pinterest, Google+, etc are a few type of Social Media Communication which is among the highly recommendable in today's era. People are becoming more interested in and relying on the social media i.e. blogs, wikis, news, online forums, for information, news, and opinion of other users on diverse subject matters.

1.2 Data Mining

Simply, data mining is outlined as a process of extracting usable information from a bigger set of any raw data and analyzing hidden patterns of data according to different perspectives for categorization into helpful information, which is collected and assembled in common areas, like information warehouses, for efficient analysis, data processing algorithms, facilitating business decision-making and alternative data needs to ultimately cut prices and increase revenue. It has applications in multiple fields, like science and research.

1.3 Social Media Analysis

Data mining provides a wide range of techniques for detection useful data from huge datasets. These techniques use data analysis, data pre-processing, and data interpretation processes within the course of data analysis. The voluminous nature of social media datasets needs automated information processing for analyzing it at an affordable time. Interestingly, data processing techniques additionally need vast knowledge sets to mine remarkable patterns from data; social network sites seem to be excellent sites to mine with data processing tools.

1.4 Social media analytics

We begin with definitions of a number of the key techniques related to analyzing unstructured textual data:

• Natural language processing—(NLP) could be a field of computer science, AI and linguistics concerned with the interactions between computers and human (natural) languages. Specifically, it is the process of a computer extracting meaningful data from language input and/or manufacturing natural language output.

• News analytics—the measuring of the assorted qualitative and quantitative attributes of matter(unstructured data) news stories. a number of these attributes are: sentiment, relevancy and novelty.

• Opinion mining—opinion mining (sentiment mining, opinion/sentiment extraction) is that the area of analysis that tries to form automatic systems to determine human opinion from text written in natural language.

• Scraping—collecting on-line information from social media and alternative websites in the kind of unstructured text and also referred to as web site scraping, internet harvesting and net data extraction.

1.5 Data Mining Techniques:

1.5.1 Classification analysis

Classification analysis allows retrieving important and significant information about the data, and the metadata. It also helps in classifying different data's in different classes. In Classification analysis algorithms is applied to decide how the new data can be classified. For example Outlook emails.

1.5.2. Association rule learning

Association rule learning allows identifying associations (dependency modeling) between different variables in the large databases. Association rule learning technique can also help in unpacking some of the hidden patterns in data which is also used to identify the variables within data and the concurrence of different variables that appear frequently in the datasets. These rules are beneficial for examining and forecasting the customer's behavior. In IT sector, association rules are used to build programs capable of machine learning.

2. LITERATURE REVIEW

[1]Zafarani R, Abbasi MA, Liu H. "Social Media Mining an Introduction. Cambridge University" ,YEAR 2014
This paper explained the social media growth as its uses increases daily basis. They explained that it is interdisciplinary field at the crossroad of disparate disciplines deeply rooted in computer science and social sciences. Researchers in this emerging field are expected to have knowledge in different areas, such as data mining, machine learning, text mining, social network analysis, and information retrieval..

[2] Subhanshu Mishra, "Analysis of Social Media Data to determine Positive and Negative Influential Nodes in the Network", year 2012
The author got tried to spot the opportunities for individuals to utilize the social media activity of their shoppers. They worked on utilization for higher service and additionally improve their own processes and explained the benefits of using calculation algorithm and a ranking algorithm to search out who the most influential positive and negativepositive influential nodes are in a very given network. In Future, the higher system can be developed which are able to predict and confirm human behaviors in a very higher manner in future.

3. PROBLEM FORMULATION

The prime objective of it is to enhance the quality of analyzed data. Social media mining is a process involving the extraction, analysis and representation of useful patterns from data in the social media, deriving from social interactions. Social media mining is a young field which has been leading research and development by handling enormous amounts of information. Just like the mining of the minerals, data mining also involve the extraction of useful information from a larger set of data, which is otherwise not evident and is difficult to acquire.

3.1 Research Methodology

Data mining provides a wide range of techniques for detecting useful knowledge from massive datasets like trends, patterns and rules. Data mining techniques are used for information retrieval, statistical modeling and machine learning. These techniques employ data pre-processing, data analysis, and data interpretation processes in the course of data analysis.

The different steps need to consider for analysis social media databases. This section provides the steps to implement the proposed work. The different steps need to follow for improve the efficiency with the algorithm.

a. Study existing techniques of data mining and analysis.

b. Research on these Techniques for identification of issues and problems.

c. enhance data mining techniques to achieve a higher level of accuracy in Sentiment analysis of social media data.

d. Implementation in MATLAB for simulation of algorithm.

e. Generate Results.

4. RESULTS

We have designed an algorithm which will mine data (comments on social media) and find occurrence of words to get more accuracy of analysis. Results after mining will be categories POSITIVE or NEGATIVE to estimate sentiments in comments. More occurred words get weight age according to number of occurrences and considered in POSITIVE or NEGATIVE category which will be decided by comparing words with list of Positive or Negative words which are stores in postitive.txt and negative.txt files.

We performed experiment with different comments. Following is result and observation.

Experiment

Data In comment.txt (tweet by Amitabh Bachan, https://twitter.com/SrBachchan/status/1001692453177643014)

Give importance to all; they that are good shall give you company, they that are bad shall give you a learning. Good bye

Run File Social's

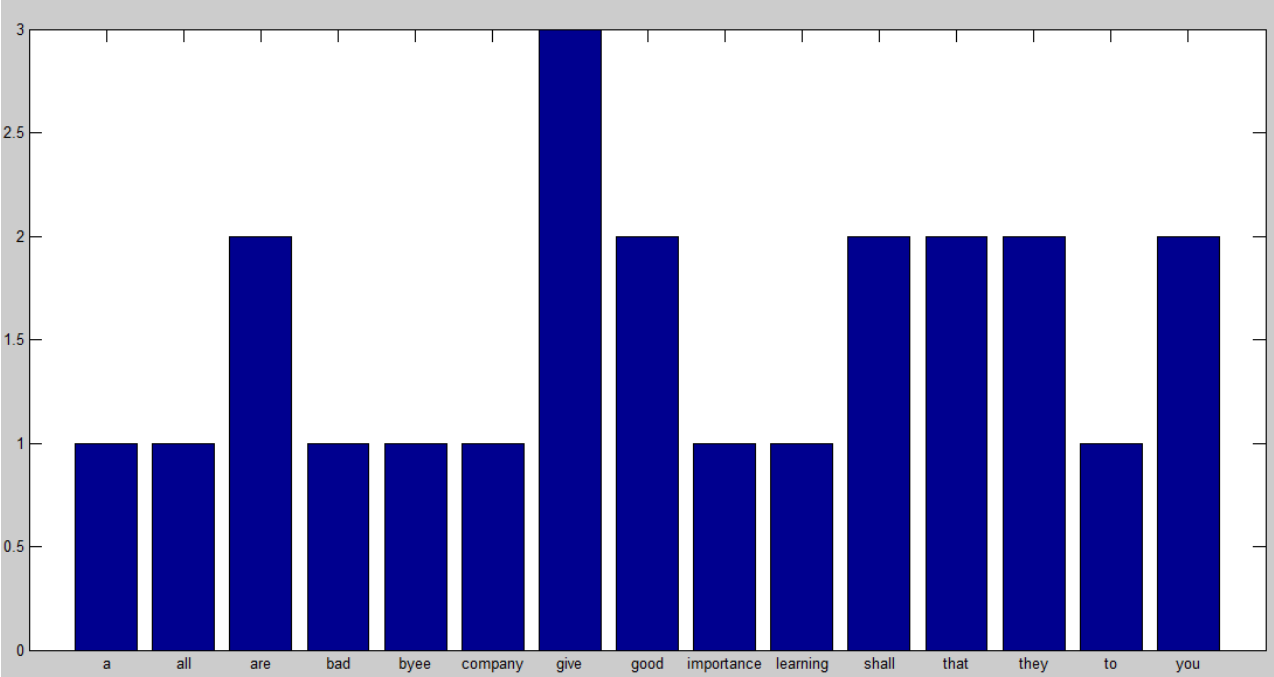

Figure 1: Data Mining Result

```
Total Positive Words 2

Pword =

    'good'    'learning'


posWeight =

     3

Total Negative Words 1

Nword =

    'bad'


negWeight =

     1

Positive Comment with 75
```

Figure 2: Sentiment Analysis

In Experiment 1, results shows that 2 words are observed as positive with total weight 3 and 1 word counted in negative list. This comment is predicted as positive comment with 75% weight age.

5. CONCLUSION

We have performed experiments by taking comments on social media and e commerce. Our analyzing tool predicted weather its positive or not accurately by comparing star rating given by user in portal. We can add more words in positive and negative text directory to make it more accurate. Further work can be done in order to increase accuracy by make dictionary stronger.

REFERENCES

[1] Mariam Adedoyin-Olowe, Mohamed Medhat Gaber and Frederic Stahl, "A Survey of Data Mining Techniques for Social Network Analysis".

[2] M. Vedanayaki, "A Study of Data Mining and Social Network Analysis".

[3] Zafarani R, Abbasi MA, Liu H. "Social Media Mining an Introduction. Cambridge University", YEAR 2014

[4] Subhanshu Mishra, "Analysis of Social Media Data to determine Positive and Negative Influential Nodes in the Network", year 2012

[5]. H.K.Chan,E.Lacka , R.W.Y. Yee and M.K.Lim "A Case Study on Mining Social Media Data"

[6] Hiroki Takikawa, Kikuko Nagayoshi "Political polarization in social media: Analysis of the "Twitter political field" in Japan"

[7] Liliya Akhtyamova, Mikhail Alexandrov, John Cardiff "Review of trends in health social media analysis"

[8] MuhammetBaykara, UğurGürtürk "Classification of social media shares using sentiment analysis"

[9] Zhaoxia Wang, Chee Seng Chong, Landy Lan, Yinping Yang, Seng Beng Ho, JooChuan Tong "Fine-grained sentiment analysis of social media with emotion sensing"

[10] Min Jing; Bryan W. Scotney; Sonya A. Coleman; Martin T. McGinnity "The Application of Social Media Image Analysis to an Emergency Management System