International Journal of Multidisciplinary Engineering in Current Research - IJMEC

Volume 10, Issue 5, May-2025, http://ijmec.com/, ISSN: 2456-4265

Opinion Mining For Social Networking Sites

Veeravalli Jinsha
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
K.Sridevi
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

Abstract:

People often express their emotions with their opinions, so by using opinions, we can know what people are thinking about any topic, such as games, politics, movies, etc. To predict sentiment from reviews, in this paper, the author is analysing people's reviews. To detect people's opinions and emotions, the author is using NLTK (natural language processing) and a data mining algorithm. For converting reviews into numeric vectors, it is called TF-IDF. We are used RF classifier for proposed methodology.

Keywords- opinion detection, data mining; sentiment analysis; data visualization; online public opinion;

I. INTRODUCTION

The continuous developments in mobile network technology have made it possible to apply new media in a way that is significantly more prevalent in the era of online big data. Social networking sites (SNS) allow users to conveniently share information online [1–2].

The positive or negative consequences of online information sharing are significantly increased when an increasing number of individuals show interest in social events, phenomena, comments, and supervision. Nonetheless, the great majority of businesses are lacking the tools necessary to quickly and accurately gather and handle competitive intelligence, information about the product market, customer attitudes, industry development, and negative information that could harm their reputation and brand value [3].

Many businesses currently use manual, traditional techniques for monitoring public sentiment. The needs of enterprises cannot be met by artificial monitoring. Although public opinion about the focal enterprise is widely dispersed online, it is difficult to accurately gather and manage. The purpose of the author's analysis in this study is to forecast sentiment from reviews.

The author is employing data mining algorithms and natural language processing (NLTK) to identify people's ideas and feelings. It is known as TF-IDF for turning reviews into numerical vectors. We employ an RF classifier in the suggested methodology.

In the proposal paper, the author uses reviews from Lending Forum, but this is not available on the internet, so we are using Amazon reviews and then coding this data mining and NLTK programme using JUPYTER, and you ask us to run the website from JUPYTER, but the website is not working from JUPYTER, so you can input reviews directly in the JUPYTER notebook and test the application for sentiment analysis. We used a Python Jupyter notebook to run the project.

This study focuses on public opinion detection through sentiment analysis and data visualisation, utilising the Natural Language Toolkit (NLTK) dataset, a valuable resource in the realm of text analysis. The primary aim of this research is to delve into the realm of sentiment analysis, leveraging the NLTK dataset to discern and analyse the sentiments expressed in textual data across diverse domains. By harnessing NLTK, an open-source library for NLP tasks, this study aims to decipher the polarity of opinions embedded within textual content, enabling a deeper understanding of public sentiment dynamics.

II. LITERATURE SURVEY

For intelligent information processing Granular computing is a new method. In a granular way it describes knowledge and simulates human thinking mode. It is appropriate for the analysis and modelling of complex and large-scale data, and it can realise the fast translation between different granular knowledge. A crucial model in granular computing is the fuzzy set, which is extensively

used in a variety of domains, including network public opinion analysis. It describes uncertain situations through fuzzy membership functions. This paper examines fuzzy sets' application research in network public opinion analysis from a big data perspective. It also addresses fuzzy sets' various roles in network public opinion analysis from four different perspectives: fuzzy reasoning, fuzzy granularity, fuzzy comprehensive evaluation, and generalised fuzzy sets. The distinctive qualities of fuzzy sets in these four areas are highlighted, and the merits of various approaches are contrasted. The particular benefits, a few issues, and the direction of this field are all covered in this paper. [1]

In service industries and modern information, one of the important technologies is social influence analysis. In social networking, big data to perform complex analysis will definitely become an important mechanism. An increasing amount of research is attracting, ranging from famous topic extraction to social influence analysis, including information diffusion modelling, influential user's identification, social influence evaluation, and the analysis and processing of big data. Based on social networking big data, we provide the architecture of social influence analysis, discuss the characteristics of social influence, and conduct a comprehensive investigation of social influence analysis. Also, social influence analysis and big data relationships are discussed. In social influence analysis based on social networking big data, these research challenges relevant to real-world issues are discussed and focus on issues in research such as effective mechanisms, evaluation metrics, network heterogeneity, causal relationships, dynamic evolution, data collection, and scalability. Our objective is to help researchers better comprehend the amount of existing work and develop new algorithms and techniques for social influence analysis by offering a comprehensive research guide for present and ongoing activities using social influence analysis in large-scale social networks. [2]

In every organisation, the analysis of social media is an essential part. To get informative insights from customer feedback about products or services, the organisation analyses the present and past records being offered. On the basis of the point of view of emotions, the analysis of the data is expressed in the dataset. For an organisation, data analysis is a treasure within social media data, where hidden information resides when the identification of expressed views is accompanied. For data analysis, which includes feedback analysis and sentiment analysis, this paper discusses the framework. In terms of complaints and suggestions to perform data analysis that not only uncovers hidden knowledge but also uncovers sentiment views, the proposed framework is an approach. By using a Naïve Bayes classifier and the N-Gramme approach over a Twitter dataset, the proposed work is performed and implemented over the Hadoop framework. By calculating precision, recall, accuracy, and F1 score, the performance of the proposed framework is examined, which is very satisfactory. [3]

Researchers from both industry and academia have welcomed the explosive growth of the social media big data mining area, which has been created by the Web's phenomenal growth in social media data. One important part of social media big data is the emotion computing of news events. Additionally, it has drawn a lot of research interest, which may help with a variety of practical uses. For example, governments may be able to monitor public opinion, and websites may be able to recommend news. Nevertheless, the majority of the sentiment computing techniques currently in use are supervised or based on conventional emotion theories, which makes them unscalable for large amounts of social media data. As such, we suggest a novel approach for performing sentiment analysis on news events. More specifically, a word emotion association network (WEAN) is constructed to jointly express the semantic and emotion of a news event based on the social media data, which forms the basis for the news event sentiment computation. To obtain the initial words emotion, a word emotion computation technique based on WEAN is proposed. These words are then improved using a standard emotion thesaurus. We are able to determine the sentiment of each statement by using the terms emotion. Experiments conducted on real-world data sets show how well the suggested approach performs when it comes to emotion computing for news events. [4]

About public opinion, the Web holds unstructured, vast, and valuable information. Along with relevant tools and techniques, the future of

opinion, current use, history, mining, and sentiment analysis are discussed. [5]

Public acceptance, information release, and distribution all affect stock market volatility. The effects of Web information on stock markets are becoming increasingly apparent as social media volume and speed increase. However, because it is difficult to automatically gather and analyse vast volumes of pertinent information, research on how Web media affects stock markets is specific and limited. In this study, 229 research publications from the domains of computer science, finance, and management information systems were thoroughly analysed in order to quantify the interaction between Web media and stock markets. Specifically, we first classified the sample works by type of medium and subsequently enumerated the fundamental methods for transforming textual data into forms that are easily interpreted by machines. we contrasted the analytic models that were employed to identify the unobserved relationships between stock movements and Web media. In order to completely comprehend the mechanisms of Web information percolation and its impact on stock markets from the perspectives of investor cognitive behaviours, corporate governance, and stock market regulation, our goal is to elucidate existing state-of-the-art research and its potential future directions. [6]

One of the major tasks of NLP is opinion mining, or sentiment analysis. In recent years, sentiment analysis has gained more attention. One of the fundamental problems of sentiment analysis is the problem of sentiment polarity categorization. This is the aim of our paper. With detailed process descriptions for sentiment polarity categorization, a general process is proposed in this paper. In this study, the data is based on reviews collected by online products from Amazon.com. With promising outcomes, the experiments for both review-level categorization and sentence-level categorization were performed. [7]

III. PROPOSED METHOD

To detect people's opinions and emotions, the author is using NLTK (natural language processing) and a data mining algorithm. In our proposal paper, we are using NLTK to read reviews datasets and process reviews, like removing stop words and special symbols, and then converting reviews into a numeric vector called TF-IDF (term frequency-inverse document frequency), and this TFIDF will be input to a data mining algorithm to train a model.

This model can be applied to any test review or opinion to predict people's emotions or sentiments. In the proposal paper, the author uses reviews from Lending Forum, but this is not available on the internet, so we are using Amazon reviews and then coding this data mining and NLTK programme using JUPYTER. You ask us to run the website from JUPYTER, but the website is not working from JUPYTER, so you can input reviews directly in the JUPYTER notebook and test the application for sentiment analysis.

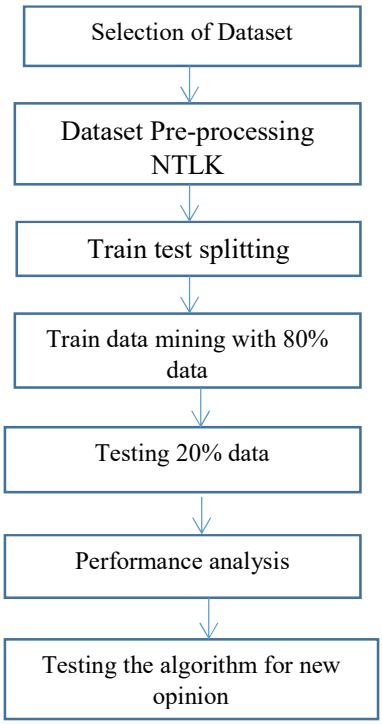

Fig.3.1 Flow chart for proposed Methodology

Figure 3.1 shows the flowchart for the proposed methodology, and the following are the steps for the proposed methodology:

1. Selection of dataset
2. Pre-processing using NLTK
3. Train test splitting
4. Training data mining with 80% data
5. Testing 20% data
6. Performance analysis

7. Testing the algorithm for new opinion

IV. RESULT

In below screen we are plotting sentiments graph so by tis visual we can understand about peoples emotion by using scores from positive and negative

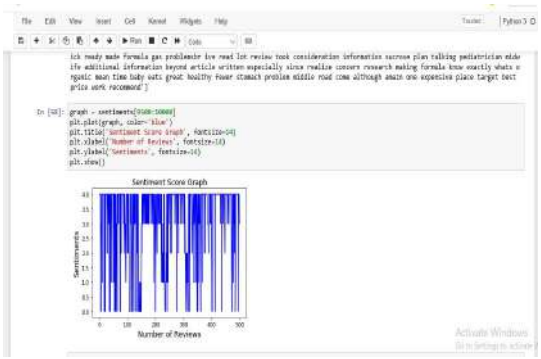

Fig.4.1 Sentiment graph

In above graph 1 to 2 are negative and then 2 to 3 neutral and 3 to 5 positive and we can see how many peoples are in positive emotion or negative and in below screen we can see same graph with dotted plot

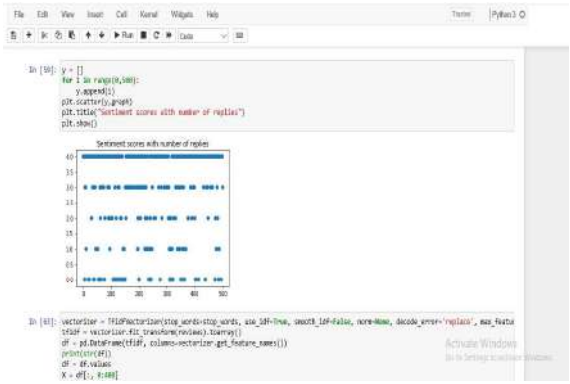

Fig.4.2 Sentiment screen with number of replies

In above graph we can see more dots are present at score 4 so more peoples emotions are positive in the dataset so by using this visual graphs we can know how many peoples are in positive or negative or neutral emotion and in below screen we are converting all reviews into TFIDF vector

In above screen all reviews words we can see top column and then review word frequency we can find below it and in below screen we are training this reviews vector with random forest data mining algorithm and calculate its prediction accuracy

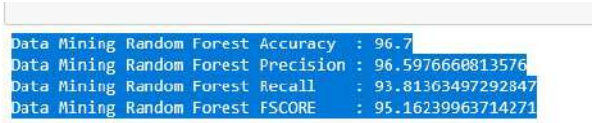

Fig.4.3 performance metrics for data mining RF classifier

In above screen with random forest we got 96% accuracy and in below screen we can enter some review and predict sentiments

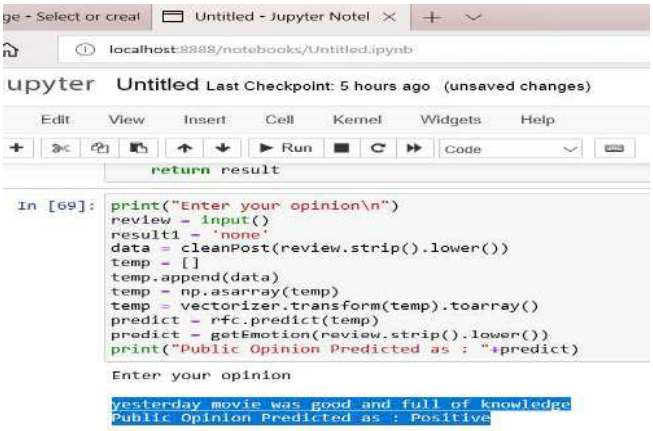

Fig.4.4 Entered opinion and its sentiment

In above screen in blue colour text first line is the review text and second line is the predicted sentiment as “positive” and now run same block to test another review

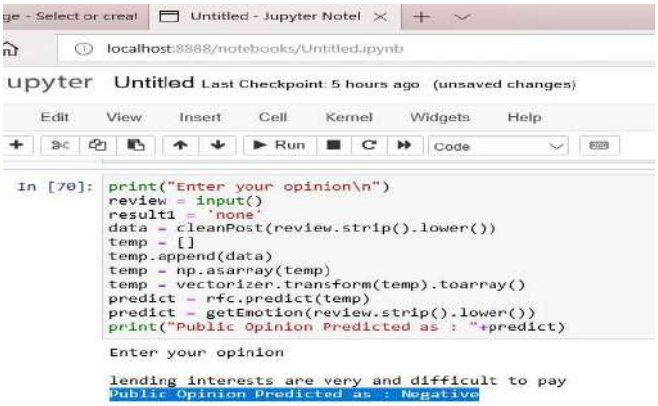

Fig. 4.5 Entered opinion and its sentiment

In above screen emotion predicted as Negative and similarly you can run that block and enter some message to predict emotion and in below screen

you can see we wrote code for flask but it not working

III. CONCLUSION

To predict sentiment from reviews, in this paper, the author is analysing people's reviews. To detect people's opinions and emotions, the author is using NLTK (natural language processing) and a data mining algorithm. We are training the review vector with a random forest data mining algorithm. We got 96% accuracy for the RF classifier.

REFERENCES

[1] J. Tang, J. Wang, and C. Wu, "Research progress on network public opinion based on rough sets from the big data perspective," Proc. IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), IEEE, May 2019, pp. 1074-1077).

[2] S. Peng, G. Wang, and D. Xie, "Social influence analysis in social networking big data: Opportunities and challenges," IEEE network, vol. 31, 2016, pp. 11-17.

[3] L. Bing and K. C. Chan, "A fuzzy logic approach for opinion mining on large scale twitter data," Proc. IEEE/ACM 7th International Conference on Utility and Cloud Computing, IEEE, December 2014, pp. 652-657.

[4] D. Jiang, X. Luo, J. Xuan, and Z. Xu, "Sentiment computing for the news event based on the social media big data," IEEE Access, vol. 5, 2016, pp. 2373-2382.

[5] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," IEEE Intelligent systems, vol. 28, 2013, pp. 15-21.

[6] Q. Li, Y. Chen, J. Wang, Y. Chen, and H. Chen, "Web media and stock markets: A survey and future directions from a big data perspective," IEEE Transactions on Knowledge and Data Engineering, vol. 30, 2017, pp. 381-399.

[7] X. Fang and J. Zhan, "Sentiment analysis using product review data," Journal of Big Data, vol. 2, 2015, p.5.

[8] S. A. Munson and P. Resnick, "Presenting diverse political opinions: how and how much," Proc. SIGCHI conference on human factors in computing systems, ACM, April 2010, pp. 1457-1466.

[9] A. Gupta and P. Kumaraguru, "Credibility ranking of tweets during high impact events," Proc. The 1st workshop on privacy and security in online social media, ACM, April 2012, p. 2.

[10] M. J. Lee and J. W. Chun, "Reading others' comments and public opinion poll results on social media: Social judgment and spiral of empowerment," Computers in Human Behavior, vol. 65, 2016, pp. 479-487.

[11] P. Sobkowicz, M. Kaschesky, and G. Bouchard, "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web," Government Information Quarterly, vol. 29, 2012, pp. 470-479.

[12] H. Chen and D. Zimbra, "AI and opinion mining," IEEE Intelligent Systems, vol. 25, 2010, pp. 74-80.

[13] X. Liu, L. Xu, and J. Zhao, "Co-extracting opinion targets and opinion words from online reviews based on the word alignment model," IEEE Transactions on knowledge and data engineering, vol. 27, 2014, pp. 636-650.

[14] C. Kaiser and F. Bodendorf, "Opinion and relationship mining in online forums," Proc. IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Vol. 1, IEEE, September 2009, pp. 128-131.

[15] D. Shen, C. Krumme, and A. Lippman, "Follow the profit or the herd? Exploring social effects in peer-to-peer lending," Proc. IEEE Second International Conference on Social Computing, IEEE, August 2010, pp. 137-144.