https://doi.org/10.46335/IJIES.2024.9.7.11 **e-ISSN: 2456-3463**
Vol. 9, No. 7, 2024, PP. 51-55

International Journal of Innovations in Engineering and Science, www.ijies.net

The Impact of Sentiment Analysis on News Articles

Dr. Vishal Tiwari, Rahel Gonjari, Ankur Singh, Janhvi Badwaik, Chaitanya Nawghare

Department of Information Technology, St. Vincent Pallotti College of Engineering & Technology, Rashtrasant Tukadoji Maharaj University, Nagpur, India

vtiwari@stvincentngp.edu.in

***Received on**: 5 May,2024* ***Revised on**: 30 June,2024* ***Published on**: 03 July ,2024*

Abstract— *The increasing importance of sentiment analysis has prompted the exploration of automated methods for tracking public perceptions of several entities. Traditionally, manual efforts were employed, but the digital age has enabled the automatic detection of sentiments from online news sources and social media platforms. This shift has motivated the entities to adopt fine-grained sentiment analysis to accurately gauge public sentiment. However, the diverse and often implicit nature of opinions in news articles presents a challenge for sentiment analysis. This study aims to address this challenge by applying a method for fine-grained sentiment analysis of news articles. The research design involves classifying sentences as positive, negative, or neutral in the context of news articles. The findings of this study have implications for marketing, public policy, and the understanding of public sentiment towards various entities.*

Keywords— sentiment analysis, news articles, public perception, public policy, digital platforms.

I. INTRODUCTION

Sentiment analysis is useful in a variety of contexts, including marketing and public policy. Its application ranges from targeted marketing to analyzing opinions about business, sports, entertainment, etc.

Several bodies, like corporations and organizations, are intensely concerned about public opinion. Traditionally, tracking reputation required human work, with individuals reading newspapers and compiling lists of favorable, negative, and neutral references. This strategy was resource-intensive and frequently entailed expensive surveys with uncertain validity. However, in the digital age, many newspapers now offer online editions.

Amidst the vast sea of information, it becomes increasingly challenging to navigate and discern the sentiments embedded within these articles. Sentiment analysis, a branch of natural language processing (NLP), emerges as a powerful tool to unravel the nuanced emotions and attitudes expressed within news content.

One of the key objectives of sentiment analysis in news articles is to explore public perception and understand how different opinions are portrayed in media coverage. By analyzing the sentiment expressed in news articles, researchers can identify trends in public. For example, sentiment analysis can reveal whether certain entities or leaders are consistently portrayed in a positive or negative light, and if there are any biases in media coverage towards specific entity. Moreover, sentiment analysis can provide insights into the overall tone and emotions associated with news articles. For example, researchers may find that articles discussing economic policies evoke more positive sentiments compared to articles discussing social issues.

Aside from traditional news sources, different emotions are routinely expressed in opinionated writings on blogs and other social media platforms. This move to digital platforms allows for the automatic detection of favorable or negative mentions of some entities in web articles, considerably lowering the labor necessary to collect such data. As a result, entities are increasingly interested in applying fine-

https://doi.org/10.46335/IJIES.2024.9.7.11 **e-ISSN: 2456-3463**
Vol. 9, No. 7, 2024, PP. 51-55

grained sentiment analysis to news stories in order to precisely grasp public opinion.

The difficulty of performing fine-grained sentiment analysis stems from the variety of ways in which the beliefs are communicated. News items, in particular, present a larger problem because they frequently omit explicit markers of sentiments. Despite their apparent impartiality, news items can transmit polarity based on the objectively positive or negative events they cover. While many sentiment analysis tools use naive keyword-based approaches, we use a similar methodology to perform fine-grained sentiment analysis, categorizing words as good, negative, or neutral in the context of different news articles.

II. LITERATURE SURVEY

In today's digital landscape, there's a pressing need to streamline the processing of unstructured data by harnessing machine capabilities. Computers excel in swiftly handling such data, necessitating efforts in mastering machine learning techniques. With the proliferation of online content such as blogs, reviews, and articles, social media now commands a significant portion of web activity. Users not only consume but actively contribute to this content, spanning various platforms and formats like articles, blogs, tweets, and multimedia uploads. Consequently, the web hosts a vast volume of unstructured textual data. The primary objective lies in sentiment analysis of this data, a burgeoning research field of considerable market interest and value. Emphasizing machine learning, this approach aims to delegate the laborious task of data processing to computers, capitalizing on their efficiency. This shift recognizes users as both consumers and creators, underscoring the dynamic nature of online participation. By leveraging machine capabilities, sentiment analysis endeavors to distill insights from the deluge of user-generated content, offering valuable perspectives in today's digital era.

A model is trained [1] on a specific dataset to classify new data. However, if the new dataset introduces novel features, as seen with the rise of web and social media data, the model's accuracy suffers as it struggles to classify documents accurately. To rectify this, there's a need to reconstruct the model using fresh training data that incorporates these new features and then assess its accuracy. Yet, this approach isn't practical given the diverse range of topics found within news corpora sourced from various outlets. Thus, the primary challenge lies in devising a universal heuristic algorithm capable of accurately discerning sentiments from these mixed news corpora. This algorithm must effectively handle document classification and feature extraction to accommodate the varied discussions and topics inherent in news articles and editorial content.

The literature survey for "Sentiment Analysis of News Tweets" [2] by Haya Fathim, published on December 23, 2021, encompasses foundational studies in sentiment analysis. Pang and Lee's work introduces sentiment analysis methodologies applicable to tweets. Liu's comprehensive survey provides insights into sentiment analysis techniques relevant for social media data like tweets. Barbosa and Feng's study on sentiment analysis in tweets offers methodologies for analyzing sentiments in short texts. Additionally, studies by Tumasjan et al. and Mohammad et al. provide insights into handling contextual influences and detecting contextual polarity in tweets. Integrating these insights forms a robust framework for analyzing sentiments in news tweets, contributing to a deeper understanding of public opinion on social media platforms.

"Sentiment Analysis of News Headlines for Stock Trend Prediction" [5] aims to bridge the gap between non-quantifiable data, specifically textual information from financial news articles, and the prediction of future stock trends through news sentiment classification. Acknowledging the influence of news on stock market volatility, the research employs text mining and sentiment analysis techniques, including Naive Bayes and Random Forest algorithms, to analyze the sentiment of news articles. As the volume of news and sources continues to escalate, the project underscores the challenge for investors to extract relevant information. Computational methods are proposed to automatically extract, filter, and aggregate news, providing real-time sentiment analysis. Emphasizing the significance of stock markets in fast-emerging economies like India, where capital infusion relies heavily on shares, the research contends that stock market prediction techniques play a crucial role in enhancing investor participation. The study recognizes the intricate relationship between market performance and the information conveyed through news, which, due to the ubiquity of internet access, constantly shapes investor sentiments. Through supervised machine learning and various text mining approaches, the research investigates the impact of news articles on stock prices, utilizing over five years of past data. The classification of news polarity and the ability to classify unknown news are key objectives, with diverse classification algorithms implemented to enhance accuracy. Ultimately, the research aims to uncover insights into the dynamic relationship between news sentiment and stock price trends, contributing valuable knowledge to the realm of stock market analysis.

https://doi.org/10.46335/IJIES.2024.9.7.11 **e-ISSN: 2456-3463**
Vol. 9, No. 7, 2024, PP. 51-55

The literature survey for [10] sentiment analysis of political communication: combining a dictionary approach with crowd coding, which introduces sentiment analysis methodologies. Wilson et al. address challenges related to negation, crucial for nuanced sentiment capture. Liu's comprehensive survey outlines sentiment analysis techniques, essential for analyzing political discourse. Works by Tumasjan et al. and Barbosa and Feng offer insights into handling complex textual structures and contextual influences. Mohammad et al.'s research on contextual polarity detection and Guerini et al.'s favorability capture methodologies are relevant for political sentiment analysis. Ma et al. and Choi et al. suggest aspect-based and fine-grained structural analysis for deeper insights. Wiebe et al.'s study on sentiment and subjectivity underscores the need for objectivity. Integrating these insights forms a robust framework for analyzing sentiments in political news articles.

SENTIWORDNET provides three numerical scores,[13] Obj(s), Pos(s), and Neg(s), for every WORDNET synset, representing objectivity, positivity, and negativity respectively. Developed via quantitative analysis of synset glosses, it employs vector term representations for semi-supervised synset classification. These scores are computed by consolidating outputs from eight ternary classifiers, each with comparable accuracy yet unique classification approaches. SENTIWORDNET is openly accessible for research and offers a user-friendly web interface.

III. METHODOLOGY

Sentiment Analysis of News articles can be done using Lexicon based approach. It can be done using supervised or unsupervised learning techniques but using supervised learning methods is a better option. In this the training data will be well labeled and will help in training any classifier model. Further this classifier will try to detect the classes on the test data. The classification will be done by analyzing the polarity of words.

This methodology is divided into six steps. Firstly data collection is to be held from different online newspapers by downloading individual news articles. Then the collected data needs to get validated. After validation, the crucial step is data annotation in which different regions of the images should be annotated. After that a model for object detection should be trained to rectify the regions automatically. Then comes Optical Character Recognition that will recognize the text from specific regions which will then get analyzed using the polarity of Words.

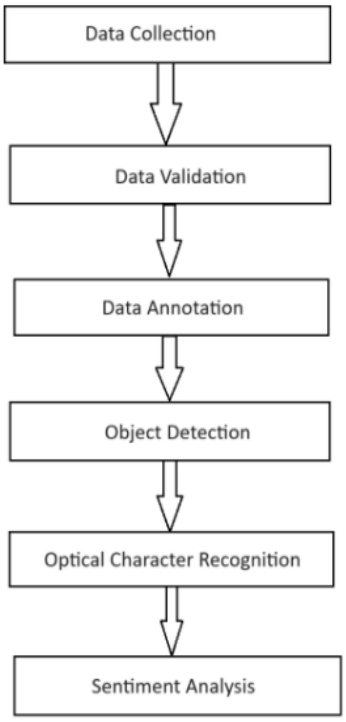


Fig. 3.1 Project pipeline

i. Data Collection

The data has been collected from various official websites of the different newspapers. From those online newspapers, the news articles have to be downloaded as an image in the .jpg extension.

ii. Data Validation

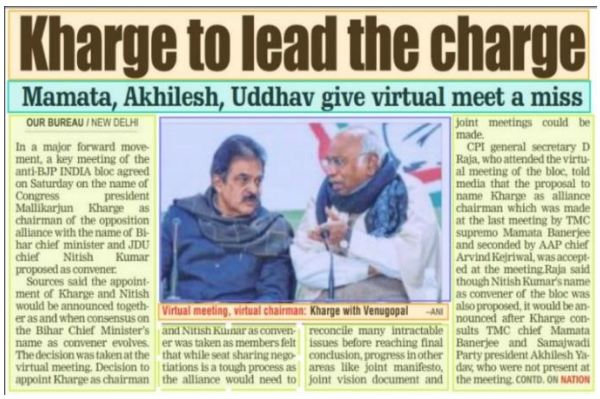
After collection of images data validation is a mandatory step. For any image to be used the main parameter is that image should not be blurred. Thus to identify such images Laplacian filter is used. The images that came out to be blur were removed from the dataset.

iii. Data Annotation

After validating the data we need to identify and label different regions of news articles such as heading, subheading, crux, content, etc. which is called as data annotation. Annotating images would make it easier to detect text of different columns of the news. This can be done using different tools such as CVAT.AI, makesense, etc.

https://doi.org/10.46335/IJIES.2024.9.7.11 **e-ISSN: 2456-3463**
Vol. 9, No. 7, 2024, PP. 51-55

International Journal of Innovations in Engineering and Science, www.ijies.net



Kharge to lead the charge

Mamata, Akhilesh, Uddhav give virtual meet a miss

OUR BUREAU / NEW DELHI

In a major forward movement, a key meeting of the anti-BJP INDIA bloc agreed on Saturday on the name of Congress president Mallikarjun Kharge as chairman of the opposition alliance with the name of Bihar chief minister and JDU chief Nitish Kumar proposed as convener.

Sources said the appointment of Kharge and Nitish would be announced together as and when consensus on the Bihar Chief Minister's name as convener evolves. The decision was taken at the virtual meeting. Decision to appoint Kharge as chairman and Nitish Kumar as convener was taken as members felt that while seat sharing negotiations is a tough process as the alliance would need to reconcile many intractable issues before reaching final conclusion, progress in other areas like joint manifesto, joint vision document and joint meetings could be made.

CPI general secretary D Raja, who attended the virtual meeting of the bloc, told media that the proposal to name Kharge as alliance chairman which was made at the last meeting by TMC supremo Mamata Banerjee and seconded by AAP chief Arvind Kejriwal, was accepted at the meeting.Raja said though Nitish Kumar's name as convener of the bloc was also proposed, it would be announced after Kharge consults TMC chief Mamata Banerjee and Samajwadi Party president Akhilesh Yadav, who were not present at the meeting. CONTD. ON NATION

Virtual meeting, virtual chairman: Kharge with Venugopal –ANI

Fig. 3.2 Annotated Image

iv. Object Detection

In this step the annotated images and its labels generated are used to train a predefined object detection algorithm "yolov8". Using this algorithm a model is trained that identifies some specific regions and creates a rectangular bounding box around it. To input this to the next step, coordinates of the bounding boxes are required. So to get these a normal python script can be used.

v. Optical Character Recognition

After the coordinates are generated, it is given as input for OCR(Optical Character Recognition) along with the image. Then different regions will be identified and text within those regions will be recognized. The text generated will be stored in text files using file handling.

vi. Sentiment Analysis

The generated text file will act as input to this step. Then pre-processing would be performed on the text i.e tokenization, lemmatization, removing stop words, etc. After this using word polarity the news will be classified as Positive, Negative or Neutral.

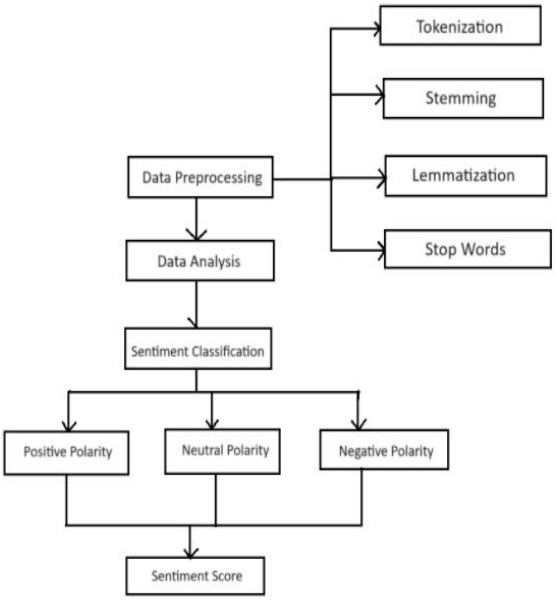


Fig. 3.3 Sentiment Analysis

IV. CONCLUSION

In recent years news has increased drastically and to analyze such a huge amount of data it would be a tedious task. The person appointed for analyzing the news would be overburdened. Thus the solution to the above problem would be to develop a system that would analyze the news articles effectively on its own. Sentiment analysis is a useful technique for figuring out how the public feels about sports, entertainment, sports and public policy. It can be used for everything from targeted advertising to spotting possibly antisocial conduct. It can assist you in understanding the most relevant and impactful feedback from your audience, which you can use to create a feedback loop.

V. FUTURE SCOPE

In the future, the field of sentiment analysis of news articles is likely to expand and evolve. Researchers can explore innovative methodologies and techniques to improve the accuracy and reliability of sentiment analysis. They can also investigate the impact of different factors, such as news context or individual characteristics, on sentiment analysis. Furthermore, as the digital landscape continues to evolve, there is a need to adapt sentiment analysis techniques to analyze sentiment in emerging platforms such as social media or online forums.

Overall, sentiment analysis of news articles is a valuable tool in understanding public opinion and assessing the impact of news on society. In addition to considering news context, researchers should also explore the impact of individual characteristics and the strength of partisan identities on sentiment analysis. This research paper aims to explore the field of sentiment analysis of news articles and its significance in understanding public perception and the impact of news.

REFERENCES

[1] *Ringki Das and Thoudam Doren Singh, "Multimodal Sentiment Analysis: A Survey of Methods, Trends and Challenges". ACM Computer Survey, 2023.*

[2] *Haya Fathim, "Sentiment Analysis of News Tweets", Rochester Institute of Technology, 2021*

[3] *Zad, Samira, Maryam Heidari. "A survey on concept-level sentiment analysis techniques of textual data." 2021 IEEE World AI IoT Congress (AIIoT). IEEE, 2021.*

[4] *Alonso, M.A.; Vilares, D.; Gómez-Rodríguez, C.; Vilares, J. Sentiment Analysis for Fake News Detection. Electronics, 2021.*

[5] *Oshi Gupta, "Sentiment Analysis of News Headlines For Stock Trend Prediction", IJRTI ,ISSN: 2456-3315 ,Volume 5, Issue 12 , 2020*

[6] *Drus, Zulfadzli, and Haliyana Khalid. "Sentiment analysis in social media and its application: Systematic literature review." Procedia Computer Science, 2019.*

[7] *P. Karthika, R. Murugeswari and R. Manoranjitham, "Sentiment Analysis of Social Media Network Using Random Forest Algorithm," IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), 2019*

[8] *M. U. Islam, F. B. Ashraf, A. I. Abir and M. A. Mottalib, "Polarity detection of online news articles based on sentence structure and dynamic dictionary," 2017 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, 2017, pp. 1-5. doi: 10.1109/ICCITECHN.2017.8281777*

[9] *V. S. Shirsat, R. S. Jagdale and S. N. Deshmukh, "Document Level Sentiment Analysis from News Articles," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, 2017, pp. 1-4.*

[10] *Aruna Gunda and Varsha Teratipally, "Sentiment Analysis of Political News articles and the effect of negation scope" International Research Journal of Engineering and Technology(IRJET), Volume:03, Issue:10, 2016*

[11] *M. Devika, C. Sunitha, and A. Ganesh, "Sentiment analysis: a comparative study on different approaches," Procedia Computer Science, vol. 87, pp. 44–49, 2016.*

[12] *V. Kharde, and P. Sonawane, "Sentiment analysis of twitter data: a survey of techniques," International Journal of Computer Applications, vol. 139, no. 11, pp. 5–15, 2016.*

[13] *A. Agarwal, V. Sharma, G. Sikka and R. Dhir, "Opinion mining of news headlines using SentiWordNet," 2016 Symposium on Colossal Data Analysis and Networking (CDAN), Indore, 2016, pp.1-5.doi: 10.1109/CDAN.2016.7570949*

[14] *Z. Madhoushi, A. R. Hamdan, and S. Zainudin, "Sentiment analysis techniques in recent works," 2015 Science and Information Conference (SAI), 2015*

[15] *Ubale Swati, Chilekar Pranali, Sonkamble Pragati "Sentiment Analysis of news articles using machine learning approach", International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835,Volume-2, Issue-4, April-2015.*