# Sentiment Analysis and Sarcasm Detection using Deep Learning

**[1]Pandarkar Abhijeet Chandrakant, [2]Mokashi Shraddha Sudhir, [3]Garje Amol Kailas, [4]Shelake Swati Bharat**

Dept. Computer Engineering
VPKBIET
Baramati, India.

***Abstract*- Interactive online environments have become cru- cial outlets for expressing opinions and sentiments. Sentiment analysis, categorizing thoughts as positive, negative, or neutral, aids in areas such as marketing and behavior analysis. Often, sentiments are conveyed with sarcasm, posing a challenge for conventional sentiment classifiers. Current research treats sen- timent and sarcasm as distinct tasks, but our work suggests they are interconnected. We propose a deep neural network to improve sentiment analysis by considering the correlation with sarcasm. Our approach outperforms existing methods, achieving a 94 percent F1-score, highlighting the effectiveness of jointly addressing sentiment and sarcasm.**

***Keywords:* Sentiment analysis, Sarcasm detection, Deep learning algorithm, Natural Lanaguage Processing. Multi-task learning.**

## I. INTRODUCTION

In today’s digital age, nearly 63 percent of the global pop- ulation, approximately 5.03 billion people, are active internet users [1], with 93 percent engaged in social media. Platforms like Twitter, Reddit, and Facebook have become integral to our daily lives, serving as spaces to share everything from personal opinions to business events. [1] Social media is not just a means of personal expression; it also serves as a real- time information hub, influencing people socially, politically, and economically.

Businesses leverage social media to connect directly with consumers, shaping their strategies based on customer feed- back and reviews. A single product review can impact con- sumer decisions, with 93 percent of internet users influenced by customer reviews in their purchases [2] [5] according to a survey by Podium. Social media’s impact extends to critical events, as seen during the COVID-19 crisis [4] where it became a platform for expressing emotions and sentiments. Analyzing these expressions becomes vital [16] sentiment analysis, a technique in natural language processing (NLP), plays a key role.

Sentiment analysis involves [2] [10] identifying and classi- fying subjective information in unstructured text to determine the polarity of sentences. It proves essential for extracting valuable insights from unstructured data sources like tweets and reviews. The importance of sentiment analysis is further highlighted in its role during the COVID-19 pandemic [4] helping the government understand public concerns and take appropriate measures.

However [7] automated sentiment analysis has its limita- tions, especially when faced with the complexity of natural language and the uncertainty of posted content. [16] The study of tweets, for instance, can be challenging due to the inclusion of hashtags, emoticons, and links, making sentiment iden- tification difficult Sarcasm poses an additional challenge as people often use positive words to express negative sentiments sarcastically, making it tricky for sentiment analysis models.

This paper addresses the impact of sarcasm detection on sentiment analysis, aiming to enhance existing models for improved accuracy and intelligent information extraction. The proposed framework involves deep multi-task learning, train- ing models for sentiment analysis and sarcasm detection si- multaneously. This approach reduces complexity and increases efficiency, offering a more intelligent and accurate way to extract information from complex textual data.

## II. LITERATURE SURVEY

Twitter sentiment analysis has attracted considerable atten- tion from researchers in recent years, resulting in a variety of new projects centered on tweet classification.However, the type of classification and the variables utilized change significantly relying on the specific objectives of the research. Yik Yang Tan [15] To enhance sentiment analysis and sarcasm detec- tion on social media, the researchers used a deep multi-task learning model using a neural network. The proposed method results in 94% accuracy, over previous methods by 3%.The major advance is in discovering the relationship between sarcasm identification and sentiment analysis, which addresses the problem of misunderstanding sarcastic statements. The study emphasizes the effectiveness of multi-task learning in enhancing sentiment analysis accuracy in social media text classification.

The research introduces the first English-Hindi code-mixed dataset for sarcasm detection in tweets. [10] conducted by Sahil Swami and his colleagues at the Language Technologies Research Centre, International Institute of Information Tech- nology, Hyderabad, addresses the increasing significance of sarcasm detection in social media. The authors use a Random Forest classifier and 10-fold cross-validation, achieving an impressive average F-score of 78.4%. They consider features such as word and character n-grams, sarcasm indicative to- kens, and emoticons in their classification system. The study highlights the dataset's potential for training and evaluating sarcasm detection systems in code-mixed tweets, emphasizing its relevance in multilingual social media analysis.

The authors [13]suggest an innovative method for identify- ing sarcasm in textual information, focusing on the challenges posed by the dynamic and vast nature of information on social media platforms. They introduce a deep neural network model consisting of a Bidirectional LSTM layer whose hyperparam- eters are optimized using a evoluting appraoch, followed by a Convolutional Neural Network (CNN) for enhanced sarcasm detection. The algorithm employs BERT as an embedding layer to capture contextual relationships, and the LSTM layer is fine-tuned through genetic optimization. The CNN layer further refines the framework's capabality to detect sarcasm. The authors discuss the importance of contextual data and claim that their proposed model is expected to outperform existing models, providing an accuracy range of 93-95%. The utilization of genetic optimization, BERT embedding, and the combination of LSTM and CNN layers contribute to the model's robustness in handling diverse and dynamic data.

M. Bouazizi et al. [14] They utilized Part-of-Speech (POS) tags to identify and extract patterns that depict the degree of sarcasm in tweets. The pattern they devised demonstrated favorable outcomes, showcasing good results. They introduced four sets of criteria corresponding to the four types of sarcasm they outlined, employing these criteria to assess if a tweet conveys irony. The accuracy achieved was reported as 83.1%.

Lu Ren [9] The study proposes MMNSS, a Multi-Level Memory Network based on Sentiment Semantics, for im- proved sarcasm detection in sentiment analysis. Unlike exist- ing models, MMNSS considers sentiment semantics crucial for accurate detection. The first-level memory network captures sentiment semantics, and the second-level network contrasts them with sentence words. An enhanced Convolutional Neural Network addresses local information absence. MMNSS shows state-of-the-art performance on the Internet Argument Corpus and Twitter dataset, highlighting the significance of sentiment semantics in complex linguistic expressions for effective sar- casm detection.

TABLE I- COMPARISON OF DIFFERENT APPROACHES OF DEEP LEARNING OF RECENT PAPERS

| **Author** | **Dataset Used** | **Model Used** | **Results/Findings** |
|---|---|---|---|
| Yik Yang Tan., 2023 [15] | Twitter | Bi-LSTM | F score-91% for Sentiment and 92% for Sarcasm Using Mutitask Learning with Bi-LSTM F1-score 94% and 93% respectively. |
| Sahil Swami ., 2018 [10] | English-Hindi code mixed | Random Forest and SVM | F-Score 78.4% |
| Adarsh M J., 2019 [13] | Text | BERT and LSTM also CNN layer used for model | Accuracy-93.95% |

| | | | |
|---|---|---|---|
| | | robustness. | |
| M. Bouazizi et al., 2016 [14] | Twitter | Pattern Based Approach | Accuracy – 83% |
| Razali et al., 2021 [12] | Twitter | SVM, KNN, LR, DT | Accuracy – 94% (highest achieved by Logistic Re-gression) |
| Lu Ren.,2020 [9] | Text | BERT and other ML, and DL ap- proaches | The error rate of the F score re-duces by 12.9% on using multi-modal data |
| M.A.Darkunde., 2020 [16] | Amazon,Twitter | LSTM, CNN | CNN Accuracy – 85.88% ,LSTM-CNN Accuracy – 93% to 94% |
| Cai et al., 2019 [26] | Twitter (image, text) | Bi-LSTM | F score- 83.44% (text), F score- 80.18% (images), |
| Baruah et al., 2020 [27] | Twitter and Red-dit | BERT, LSTM, BiLSTM | F-score (BERT)- 0.743(Twitter), 0.658(Reddit) |
| Srivastava et al., 2020 [28] | Twitter and Red-dit | Hierarchical BERT | F score 0.74(Twitter), 0.639(Reddit) |

## III. METHODOLOGY

This paper introduces a model for sentiment analysis that takes into account potential existence of sarcasm.

### *A. Data Gathering*

Data gathering serves as the initial phase in machine and deep learning, often requiring substantial time and resources to collect potentially relevant data. The proposed framework relies on two essential datasets: a sentiment dataset and a sar- casm dataset.Twitter data set ,News Headline Dataset and other dataset like twitter and Raddit [23]Other popular datasets for recognizing sarcasm include Amazon and Facebook datasets.

### *B. Data Pre-Processing*

Datasets sourced online or gathered manually are often disorderly and lack structure, requiring pre-processing to en-sure accuracy during training. [15] This paper employs a consistent pre-processing approach for both sentiment and sar- casm datasets, leveraging shared natural language processing (NLP) techniques. The pre-processing steps include removing irrelevant words like hyperlinks, noisy terms (e.g., retweets, Precision is the accuracy of positive predictions, and it is defined as the ratio of true positives to the total predicted positives, including both true positive (TP) and false positive (FP) as given in Eq. 2.

TABLE II- THE DIFFERENT DATASETS USED BY RECENT RESEARCHERS.

| **Sr No.** | **Referred Paper** | **Dataset Name/ Extracted From** | **Type of data** | **Description** |
|---|---|---|---|---|
| 1 | [20] | Facebook | Text | 3000 comments |
| 2 | [21] | Twitter | Text | Total(980)<br>502 (neutral),<br>250 (positive),<br>228 (negative) |
| 3 | [22] | Twitter | Text | 6000 tweets |
| 4 | [23] | Twitter and Reddit | Text | Twitter 163K Tweets,<br>Reddit 37K Comments |
| 5 | [24] | News Head-line | Text | - |

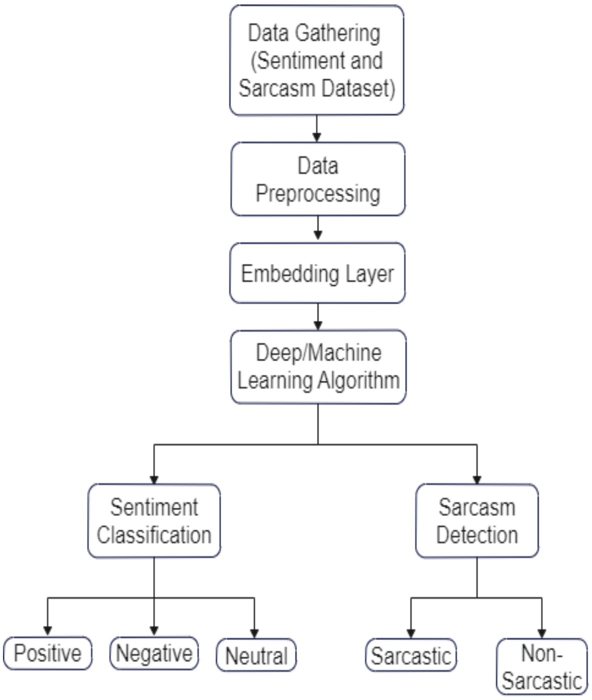


Fig. 1. Architecture

stock market tickers), and stop words. Punctuations are also eliminated, and word stemming is applied to reduce words to their base form. Tokenization is then employed to convert text into a vector representation suitable for input into the deep learning model, streamlining the datasets for effective analysis.

*C. Neural Newtwork*

Numerous classifiers and rule-based strategies are utilized to tackle sarcasm detection as a binary classification task. Figure
2 illustrates various methodologies employed for sentiment classification and sarcasm detection. [25] techniques used for sentiment classification and sarcasm detection:

1) Support Vector Machine (SVM)

SVM technology excels in sentiment analysis [29] [30] by handling binary and multi-score tasks, integrating with EWGA for optimization, evaluating feature sets, analyzing stylistic features, optimizing feature weights, and validating models with cross-validation, ensuring accurate sentiment analysis.

2) Na¨ıve Bayes (NB)

In their study on sentiment analysis and sarcasm detection in Indonesian social media, the authors utilize the Na¨ıve Bayes algorithm within a framework comprising pre-processing, feature extraction, and classification components. Through pre-processing, informal language is standardized, and features like unigrams and sentiment scores from SentiWordNet are extracted. [31] Na¨ıve Bayes is employed for sentiment classification and sarcasm detection, leveraging these features alongside additional ones such as negativity and interjection words. Experimental results suggest that incorporating sentiment scores improves sentiment analysis accuracy, while the inclusion of features like negativity and interjection words enhances sarcasm detection accuracy when using Na¨ıve Bayes.

3) Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) represents a specialized version of recurrent neural network (RNN) architecture, strategically crafted to tackle the issues encountered by conventional RNNs. One notable challenge addressed by LSTM is the vanishing gradient problem, a hindrance to the conventional RNNs' capacity to effectively learn and preserve long-term dependencies within sequential data. LSTM networks are characterized by memory cells that incorporate diverse gates, such as input, forget, and output gates. These gates play a crucial role in controlling the information flow within the network. Through these gates, LSTMs can effectively choose to retain or discard specific information over time. This selective mechanism enables LSTMs to acquire and remember dependencies across extended sequences. [15] In natural language processing (NLP), LSTM networks have evolved into a powerful tool for tasks involving sequential data, such as sarcasm detection and sentiment analysis. By capturing the relationships between words in a sentence and preserving context from preceding parts of the text, LSTMs excel at understanding the overall meaning and sentiment conveyed by language. In sentiment analysis, for instance, LSTM networks can effectively discern the emotional tone of a given text whether neutral, positive, or negative, by analyzing the sequence of words and their respective contexts. Moreover, in sarcasm detection, LSTMs are adept at recognizing subtle linguistic cues and understanding the underlying intentions behind seemingly contradictory statements. This ability to model long-term dependencies makes LSTM networks well-suited for a vast range of NLP tasks, where understanding the sequential nature of language is paramount to achieving

3) F1-score is a helpful metric to compare two classifiers. F1 score considers both recall and precision, which is defined as Eq. 3.

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

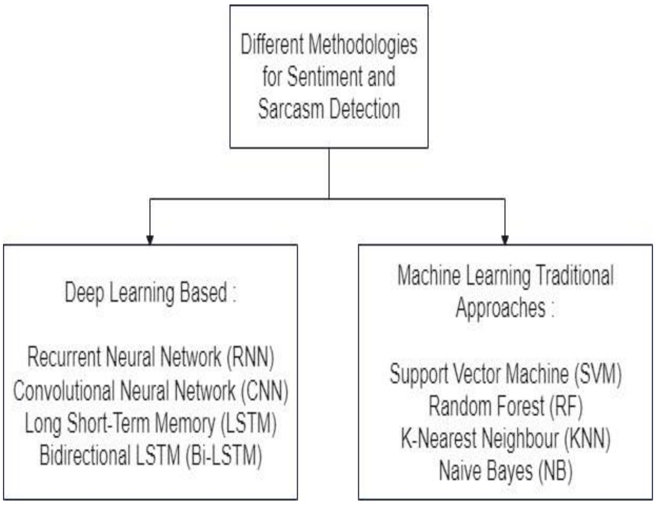


Fig. 2. Different Methodologies for Sentiment and Sarcasm Detection

4) Accuracy of a model can be calculated using the fol- lowing formula, which is defined as Eq. 4
5)

zAccuracy = $TP + TN$
$TP + FP + FN + TN$

TABLE III- APPROACH USED IN REFERRED PAPER

| Approach | Referred paper |
|---|---|
| SVM | [10], [12], [14], [29], [30] |
| NB | [31], [15], [12] |
| RNN | [15], [32] [13] |
| CNN | [13], [16] |
| LSTM/BiLSTM | [15], [27], [26], [28] |

accurate predictions and insightful analyses.
4) Recurrent Neural Network (RNN)
Recurrent Neural Networks (RNNs) are utilized in sentiment analysis and sarcasm detection by processing text sequences one word at a time, updating a hidden state at each step. [15] For sentiment analysis, the final hidden state is passed through a fully connected layer to predict sentiment. Similarly, for sarcasm detection, the network learns linguistic patterns to determine sarcasm, outputting a binary prediction. However, traditional RNNs struggle with long-range dependencies, which is addressed by advanced architectures like LSTM and GRU.These models excel in preserving information across extended sequences, making them preferable for tasks that demand a grasp of context and sentiment.

*D. Evaluation metrics*
The following standard metrics are used in this paper to benchmark the performance of the proposed framework.
1) Recall is a measure of the ability of a model to detect positive in each sentence (also known as sensitivity). It is the ratio of true positive (TP) to the total of True Positive and false negative (FN) as given in Eq. 1.
$TP$

## IV. CONCLUSIONS AND FUTURE WORK

In conclusion, this paper highlights the advancements in sentiment analysis and sarcasm detection. It discusses efforts using different datasets and methods, stressing the growing importance of sarcasm detection. The limitations of relying solely on textual data are noted, with emphasis on the effec- tiveness of multimodal datasets. The paper also addresses the role of memes and non-verbal cues in sarcasm communication. Moreover, it outlines the process of data retrieval, prepro- cessing, feature extraction, and classification techniques in sar- casm detection. It underscores the importance of considering word sentiment before and after classification to accurately analyze sarcastic sentence polarity. The findings stress the need for comprehensive approaches integrating various methods and background knowledge for accurate sarcastic message
identification and analysis.

## REFERENCES:

1. Statista Research Department. (2022, September 20). Internet and social media users in the world 2022. Statista. Retrieved October 6, 2022, from https://www.statista.com/statistics/617136/digitalpopulationworldwide/
2. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., and Cao, B(2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan China. The Lancet, 395(10223), 497–506. https://doi.org/10.1016/s01406736(20)30183-5
3. American Psychological Association. (n.d.). APA: U.S. adults report highest stress level since early days of the COVID-19 pandemic. Amer- ican Psychological Association. Retrieved October 6, 2022, from https: //www.apa.org/news/press/releases/2021/02/adults-stress-pandemic
Recall =$TP + FN$
4. American Psychological Association. (n.d.). APA: U.S. adults report highest stress level since early days of the COVID-19 pandemic. Amer- ican Psychological Association. Retrieved October 6, 2022, from https: //www.apa.org/news/press/releases/2021/02/adults-stress-pandemic
5. Online Reviews Stats and Insights. Podium. (n.d.). Retrieved October 6, 2022, from

https://www.podium.com/resources/ podium-state-of-online-reviews.
6. Pang, B., and Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In:

Precision = $\frac{TP}{TP + FP}$

7. Proceedings of the 42nd Annual Meeting on Association for Computa- tional Linguistics–ACL '04. https://doi.org/10.3115/1218955.1218990
8. Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal of Un- certainty, Fuzziness and Knowledge-Based Systems, 06(02), 107–116. https://doi.org/10.1142/s0218488598000094
9. Turney, P. D. (2001). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics–ACL '02. https://doi.org/10.3115/1073083.1073153
10. Ren, Lu, et al. "Sarcasm detection with sentiment semantics enhanced multi-level memory network." Neurocomputing 401 (2020): 320-326.
11. Swami, Sahil, et al. "A corpus of english-hindi code-mixed tweets for sarcasm detection." arXiv preprint arXiv:1805.11869 (2018).
12. Kolchinski, Y. Alex, and Christopher Potts. "Representing social me- dia users for sarcasm detection." arXiv preprint arXiv:1808.08470 (2018).https://arxiv.org/pdf/1808.08470.pdf
13. Razali, Md Saifullah, et al. "Sarcasm detection using deep learn- ing with contextual features." IEEE Access 9 (2021): 68609- 68618. https://www.researchgate.net/publication/351571706 Sarcasm Detection Using Deep Learning With Contextual Features
14. Adarsh, M. J., and Pushpa Ravikumar. "Sarcasm detection in text data to bring out genuine sentiments for sentimental analysis." 2019 1st international conference on advances in information technology (ICAIT). IEEE, 2019. https://ieeexplore.ieee.org/document/8987393
15. Bouazizi, Mondher, and Tomoaki Otsuki Ohtsuki. "A pattern-based approach for sarcasm detection on Twitter." IEEE Access 4 (2016): 5477-5488.https://ieeexplore.ieee.org/document/7549041
16. Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning Yik Yang Tan1 · Chee-Onn Chow1 · Jeevan Kanesan1 · Joon Huang Chuah1 ·YongLiang Lim1 https://www.researchgate.net/publication/369011470 Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning
17. Ashok, Darkunde Mayur, et al. "Sarcasm detection using genetic op- timization on LSTM with CNN." 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.
18. Zhao, J., Liu, K., and Xu, L. (2016). Sentiment analysis: Mining opinions, sentiments, and emotions.Computational Linguis- tics, 42(3), 595–598. https://direct.mit.edu/coli/article/42/3/595/1534/ Sentiment-Analysis-Mining-Opinions-Sentiments-and
19. Diana, M., and MA, G. (2014). Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. Lrec 2014 Proceedings. https://www.researchgate.net/publication/325193763 Who cares about sarcastic tweets Investigating the impact of sarcasm on sentiment analysis
20. Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning Yik Yang Tan1 · Chee-Onn Chow1 · Jeevan Kanesan1 · Joon Huang Chuah1·YongLiang Lim1 https://link.springer.com/article/ 10.1007/s11277-023-10235-4
21. Sulaiman, Mohd Suhairi Md, et al. "Natural language processing based features for sarcasm detection: An investigation using bilingual social media texts." 2017 8th International conference on information technol- ogy (ICIT). IEEE, 2017.https://ieeexplore.ieee.org/document/8079931
22. Lunando, Edwin, and Ayu Purwarianti. "Indonesian social media sen- timent analysis with sarcasm detection." 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS). IEEE, 2013.https://ieeexplore.ieee.org/document/6761575
23. Dutta, Poulami, and Chandan Kumar Bhattacharyya. "Multi-Modal Sarcasm Detection in Social Networks: A Comparative Review." 2022 6th International Conference on Computing Methodologies and Com- munication (ICCMC). IEEE, 2022 https://ieeexplore.ieee.org/document/ 9753981
24. A, C. K. (2019). Twitter and reddit sentimental analysis dataset. Kaggle. Retrieved October 7, 2022, from httpshttps://www.kaggle.com/datasets/ cosmos98/twitter-and-reddit-sentimental-analysis-dataset
25. Misra, R. (2019). News headlines dataset for sarcasm detection. Kag- gle. Retrieved October 7, 2022, fromhttps://www.kaggle.com/datasets/ rmisra/news-headlines-dataset-for-sarcasm-detection
26. Sarcasm Detection Using Deep Learning Approaches: A Reviewhttps: //www.ijrte.org/wp-content/uploads/papers/v11i6/F74760311623.pdf
27. Cai, Yitao, Huiyu Cai, and Xiaojun Wan. "Multi-modal sarcasm detec- tion in twitter with hierarchical fusion

model.” Proceedings of the 57th annual meeting of the association for computational linguistics. 2019. https://aclanthology.org/P19-1239/

28. Baruah, Arup, et al. ”Context-aware sarcasm detection using bert.” Pro- ceedings of the Second Workshop on Figurative Language Processing. 2020.https://aclanthology.org/2020.figlang-1.12/
29. Srivastava, Himani, et al. ”A novel hierarchical BERT architecture for sarcasm detection.” Proceedings of the Second Workshop on Figurative Language Processing. 2020 https://aclanthology.org/2020.figlang-1.14/
30. Sentiment classification based on supervised latent n-gram analy- is.https://dl.acm.org/doi/10.1145/2063576.2063635
31. Sentiment analysis in multiple languages: Feature selection for opinion classification in Web orums.https://dl.acm.org/doi/10.1145/1361684. 1361685
32. Indonesian social media sentiment analysis with sarcasm detection.https://ieeexplore.ieee.org/document/6761575
33. I didn’t mean what I wrote! Exploring Multimodality for Sarcasm Detection. https://ieeexplore.ieee.org/document/9206905