DEEP LEARNING FOR SARCASM IDENTIFICATION IN TEXT DATA

SRAVAN KUMAR D
Research Scholar, Department of Computer Science and Engineering (CSE), Sikkim Alpine University, Kamrang, Namchi, (Sikkim),

DR PRASADU PEDDI
Research Guide, Sikkim Alpine Univeristy, Kamrang, Namchi, (Sikkim),

DR. VENKATESH KONDAVETI
Research Co-guide, Associate Professor, Ramachandra College of Engineering-Eluru, shravankumar1248@gmail.com

Abstract
Sarcasm is a form of expression in which individuals convey positive or negative sentiments using words that contradict their actual intent. This style of communication is increasingly prevalent in news headlines and on social media, making it difficult for readers to accurately identify sarcasm. To address this challenge, it is crucial to develop intelligent systems capable of detecting sarcasm in headlines and news articles. This research paper proposes a deep learning-based model for sarcasm detection in news headlines. The model is designed to achieve three primary objectives: (1) to understand the underlying meaning of the text or headline, (2) to learn the characteristics of sarcasm, and (3) to accurately identify sarcasm in the text. Previous studies on sarcasm detection have largely relied on tweet datasets, using hashtags to distinguish between sarcastic and non-sarcastic content. However, such datasets are often noisy due to inconsistent language use and tagging. In contrast, this study leverages multiple datasets to provide a more comprehensive understanding of sarcasm in online communication. By incorporating various forms of sarcasm from the Sarcasm Corpus V2, as well as sarcastic news headlines from The Onion and HuffPost, the proposed model is intended to generalize effectively across different contexts. The model employs Long Short-Term Memory (LSTM) networks to capture temporal dependencies in text and utilizes a GlobalMaxPool1D layer for enhanced feature extraction. Evaluation on training and test datasets yielded accuracy scores of 0.982 and 0.952, respectively, demonstrating the model's strong performance in sarcasm detection.

Keywords: *sarcasm detection; sarcasm; sentiment analysis; text data;*

1 INTRODUCTION

The Oxford Language Dictionary defines sarcasm as the use of irony to mock or express contempt. It is a widely used form of expression, yet identifying sarcasm in written text remains challenging [1,2]. This difficulty arises from the ambiguity of language and the speaker's or writer's intent, which often contrasts with the literal meaning. Sarcasm is frequently employed to attract attention and to convey negative sentiments more directly.

With the rise of online platforms such as Facebook, Instagram, and Twitter, individuals frequently share opinions on topics ranging from politics to entertainment and consumer products. The complex sentence structures and nuanced expressions used in these online messages make it challenging for both humans and machines to interpret their meaning, driving significant interest in sentiment analysis and sarcasm detection within the field of machine learning [3].

Previous research on sarcasm detection has predominantly focused on Twitter data, utilizing methods ranging from rule-based approaches to advanced machine learning techniques. Rule-based methods often categorized text based on incongruity. However, these studies were limited by unclear labeling and datasets that were not explicitly designed for sarcasm detection. More recent efforts have aimed to improve detection by analyzing the distinct characteristics of sarcastic versus non-sarcastic expressions. Despite these advances, sarcasm detection remains an ongoing area of research.

Neural networks have been employed to develop sarcasm detectors for various content types, including tweets and Flickr posts. Building on this work, the present study proposes an enhanced Long Short-Term Memory (LSTM) model for sarcasm detection, referred to as GMP-LSTM. This model integrates a GlobalMaxPooling layer, enabling it to capture critical features from temporal sequences and identify higher-level patterns in the data compared to the standard LSTM.

The proposed GMP-LSTM was evaluated on two sarcasm datasets and demonstrated higher accuracy than the original LSTM model. These results indicate that the GMP-LSTM is a promising tool for detecting sarcasm in textual data.

The paper is structured as follows: Section 2 reviews prior research on machine learning-based sarcasm detection. Section 3 details the methodology, including the proposed deep learning architecture for detecting sarcasm in news headlines. Section 4 presents the implementation of the classifier and the evaluation results. Finally, Section 5 summarizes the study's findings and discusses their implications.

2. LITERATURE REVIEW

Previous research on sarcasm detection has primarily focused on Twitter datasets, although more recent studies have expanded to include written news headlines from professional journalists, such as those exclusively containing sarcastic content from The Onion. In contemporary natural language processing (NLP), convolutional neural networks (CNNs) and Long Short-Term Memory (LSTM) networks are commonly employed methodologies [4]. For example, M. S. Razali [5] proposed a multichannel attention-based Bidirectional LSTM (BLSTM) technique for detecting sarcasm in news headlines. Other approaches have applied classical machine learning methods, including Naive Bayes classifiers and support vector machines, using features such as punctuation, phrase structure, and hashtags, achieving accuracies above 75% across social media datasets.

Deep learning techniques have increasingly outperformed traditional shallow models, leading most researchers to favor automatic feature extraction over manual methods [5,6]. Early sarcasm detection efforts focused on tweets, often using contextual clues such as hashtags (#sarcasm, #sarcastic, #notsarcasm) as indicators [3,7,8]. Feature sets from Twitter datasets, such as one consisting of 9,400 tweets with 5,800 positive, 3,100 negative, and 500 neutral examples [9], were commonly used. CNNs and LSTMs have remained central to NLP-based sarcasm detection [10], while attention-based BLSTM models have been applied to news headlines [11].

Additional strategies included pattern-based rules, such as counting positive/negative words, emotional terms, or recurring sarcastic keywords like "yay!" and "nay!" [12]. The short length of tweets (historically 140 characters) poses a limitation, making it harder to detect sarcasm and understand intended meaning [15,16]. Previous work has classified tweets based on sentiment polarity, context (e.g., events, opinions, transactions, private messages), slang, and humor-related features [17,18,19]. Hashtags and historical tweet context have also been leveraged to infer tone, although such approaches depend on the availability of sufficient historical data [20].

Recent advancements combine pre-trained word embeddings with deep learning architectures. For instance, CNN-LSTM models using embeddings such as Word2Vec have achieved state-of-the-art performance [21]. Attentive RNNs, when paired with pre-trained embeddings, effectively capture contextual patterns in textual data. In one study [21], a news headline dataset in JSON format included 26,700 headlines (11,700 sarcastic from The Onion and 14,900 non-sarcastic from HuffPost). Headlines were tokenized using Keras, converted into sequences of 300 words, and reduced to 20,000 unique words. A pre-trained embedding matrix ([20,000 × 300]) from Stanford was then fed into a BLSTM model, followed by a dense neural network to output sarcasm labels.

In another study [22], pre-trained user embeddings were initially considered but later removed, as sarcasm in news headlines does not depend on the author. The LSTM encoded word sequences bidirectionally, with forward and backward passes capturing context in both directions. The dataset was split into training, validation, and test sets in an 80:10:10 ratio. This model achieved 89.7% accuracy, improving upon the baseline of 84.88% by approximately 5%, demonstrating the effectiveness of LSTM-based architectures for sarcasm detection.

3 MATERIALS AND METHODS

The methodology of this study encompasses several key components, ranging from data collection to model evaluation. The overall approach to sarcasm detection is illustrated in Figure 1. Additionally, a thorough review of the relevant literature was conducted to assess and evaluate the existing techniques for sarcasm detection.

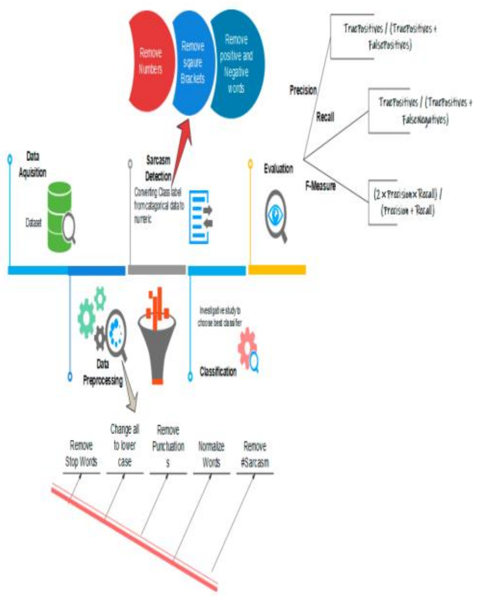

Figure 1. Methodology for sarcasm detection.

3.1 Data Acquisition

The news headlines dataset used in this study offers several advantages over typical Twitter datasets. It features formal language with minimal spelling errors, high-quality labels—thanks to The Onion's exclusive focus on satirical news—and independent headlines that are not replies to other posts. Each entry in the dataset includes three attributes: the headline text, a sarcasm label, and a link to the corresponding article [31].

The second dataset, The Sarcasm Corpus V2 [32], is a subset of the Internet Argument Corpus (IAC) and includes posts annotated for sarcasm. It covers three categories of sarcastic language: general sarcasm, hyperbole, and rhetorical questions. This dataset was selected for its diverse range of sarcastic expressions and annotated markers, making it well-suited for training and testing a sarcasm detection model.

3.2 Data Preprocessing

Data preprocessing involved cleaning and merging both datasets. The News Headlines dataset contains 26,709 entries, including 14,985 neutral headlines from HuffPost and 11,724 satirical headlines from The Onion. Preprocessing steps included removing article links and text in square brackets, converting all text to lowercase, and removing punctuation and numerals. Headlines were then sorted in ascending order based on length.

For The Sarcasm Corpus V2, no preprocessing was required since the dataset already included sarcasm annotations. Both datasets were merged, and a common label "GEN" was assigned to all headlines. Combining the datasets increased corpus diversity: the News Headlines dataset contributed formal, structured language, while The Sarcasm Corpus V2 added sarcastic and humorous expressions. This diversity improved the robustness of the sarcasm detection model during training.

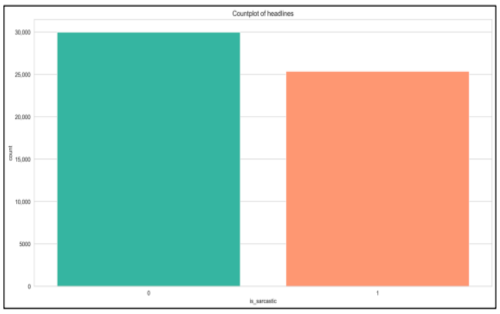

Figure 2. Row count for sarcastic and non-sarcastic data

In Figure 3, words are visualized as a word cloud, where the size of each word reflects its relative frequency of occurrence compared to the other words.

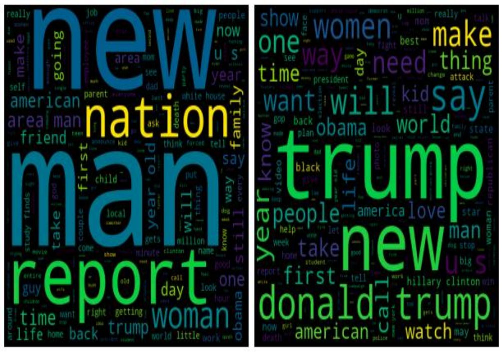

Figure 3. Word clouds of sarcastic and non-sarcastic headlines.

4 MODEL SELECTION

In this study, experimentation began with classical machine learning (ML) classifiers to detect sarcastic entries within the datasets. The classifiers evaluated included Decision Tree, Random Forest, Multinomial and Bernoulli Naive Bayes, and Support Vector Machines (SVM). The Decision Tree model is a non-parametric method that partitions the input space into regions corresponding to specific classes, while the Random Forest is an ensemble of decision trees that aggregates predictions for improved accuracy. Multinomial and Bernoulli Naive Bayes models are well-suited for multiclass and binary text classification tasks, respectively. SVM is a linear model that separates the input space using a hyperplane and has demonstrated strong performance in text classification tasks.

In addition to these classical classifiers, we implemented a standard Long Short-Term Memory (LSTM) model using a Sequential deep learning architecture. The model consisted of an embedding layer, followed by two dense layers with ReLU activation, two Dropout layers for regularization, and a final dense layer with Sigmoid activation for output. The embedding layer generated dense word representations, while the LSTM layer captured temporal dependencies within the sequences. The dense layers captured higher-level patterns in the data. Despite achieving reasonable performance (accuracies ranging from 64.87% to 84.25%), these models highlighted the need for an improved approach.

To address this limitation, we propose a novel deep learning architecture called GMP-LSTM, which integrates Global Max Pooling with LSTM networks. The experimental results, summarized in Table 1, demonstrate that the GMP-LSTM model outperforms classical classifiers, achieving accuracy above 90%. This performance indicates that the GMP-LSTM is a robust and effective approach for detecting sarcasm in textual data. The following section provides a detailed description of the GMP-LSTM model and its potential applications.

Table 1. Classifier scores.

Classifier	**Precision**	**Recall**	**F1**	**Accuracy**
Decision tree	71.2407%	70.4845%	70.8606%	74.6462%
Random forest	87.8693%	83.2642%	85.5048%	80.8162%
Multinomial	88.8000%	73.4013%	80.3697%	83.5177%
Bernoulli	84.7538%	80.1688%	82.3976%	84.2549%
Support vector	82.0578%	81.2152%	81.2215%	84.2245%
LSTM model	62.3196%	78.3671%	69.4281%	64.8793%

In the proposed model, the original LSTM layer was replaced with a GlobalMaxPool1D layer, which performs max pooling operations over temporal sequences to capture the most salient features. The architecture also includes four dense layers with ReLU activation functions, interleaved with four Dropout layers for regularization, and a final dense layer with a Sigmoid activation function to produce the output. The dense layers extract higher-level patterns from the data, enhancing the model's ability to detect sarcasm.

To further improve performance, GlobalAveragePooling and pruning techniques were incorporated. The model was trained for 100 epochs and consists of a total of 10 layers, optimizing feature extraction while maintaining robustness.

5 RESULTS

In this study, machine learning techniques were employed to detect sarcastic headlines, including classifiers such as Random Forest, Multinomial Naive Bayes, Bernoulli Naive Bayes, Decision Tree, and

Support Vector Machines. The data were preprocessed using tokenization, with the Keras tokenizer converting headlines into a suitable format for analysis. An embedding matrix was created, and a GMP-LSTM model was constructed using binary cross-entropy as the loss function and the Adam optimizer.

The performance of the GMP-LSTM model in detecting sarcastic headlines was evaluated using various metrics, including training, validation, and testing accuracy, as presented in Table 2. The GMP-LSTM model achieved a peak accuracy of 92%, outperforming Bernoulli Naive Bayes and Support Vector classifiers, demonstrating its effectiveness for sarcasm detection in news headlines.

Table 2. Performance matrix for applied classifiers.

Classifiers	**Precision**	**Recall**	**F1**	**Accuracy**
Decision tree	71.2407%	70.4845%	70.8606%	74.6462%
Random forest	87.8693%	83.2642%	85.5048%	80.8162%
Multinomial	88.8000%	73.4013%	80.3697%	83.5177%
Bernoulli	84.7538%	80.1688%	82.3976%	84.2549%
Support vector	82.0578%	81.2152%	81.2215%	84.2245%
LSTM model	62.3196%	78.3671%	69.4281%	64.8793%
Proposed GMP-LSTM	**98.2501%**	**97.5473%**	**98.3913%**	**92.5486%**

The performance of different classifiers was evaluated using the dataset, with results summarized in Table 2. The evaluation metrics included precision, recall, F1-score, and accuracy. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positives among all actual positive instances. The F1-score, the harmonic mean of precision and recall, provides an overall measure of classifier performance. Accuracy represents the proportion of correct predictions out of all predictions made.

The results indicate that the GMP-LSTM model achieved the highest accuracy of 92.55%, along with high precision, recall, and F1-scores. Other classifiers demonstrated varying levels of performance, with Decision Tree and Random Forest models showing the lowest accuracy scores of 74.65% and 80.82%, respectively. These findings highlight the GMP-LSTM model as a robust and effective approach for text classification and sentiment analysis tasks.

6 CONCLUSIONS

This research presents an effective approach for detecting sarcasm in text using the proposed GMP-LSTM model. Experimental results demonstrate that the model achieves high accuracy in distinguishing sarcastic from non-sarcastic sentences, outperforming other classifiers such as Decision Tree, Random Forest, Multinomial and Bernoulli Naive Bayes, and Support Vector Machines.

The findings underscore the potential of machine learning techniques for sarcasm detection and provide valuable insights for developing more accurate and efficient models. The proposed GMP-LSTM model can be applied to a variety of real-world tasks, including social media monitoring, sentiment analysis, and online customer service, enhancing both effectiveness and efficiency. Future research may further improve performance by exploring additional features and optimization techniques. Overall, this study contributes to advancements in natural language processing and machine learning.

REFERENCES

1. *Dave, A.D.; Desai, N.P. A comprehensive study of classification techniques for sarcasm detection on textual data. In Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 3–5 March 2016; pp. 1985–1991.*
2. *Cheang, H.S.; Pell, M.D. The sound of sarcasm. Speech Commun. 2008, 50, 366–381. [CrossRef]*
3. *Riloff, E.; Qadir, A.; Surve, P.; De Silva, L.; Gilbert, N.; Huang, R. Sarcasm as contrast between a positive sentiment and negative situation. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, DC, USA, 18–21 October 2013; pp. 704–714.*
4. *Mandal, P.K.; Mahto, R. Deep CNN-LSTM with word embeddings for news headline sarcasm detection. In Proceedings of the 16th International Conference on Information Technology-New Generations (ITNG 2019), Las Vegas, NV, USA, 1–3 April 2019; pp. 495–498.*
5. *Razali, M.S.; Halin, A.A.; Ye, L.; Doraisamy, S.; Norowi, N.M. Sarcasm Detection Using Deep Learning With Contextual Features. IEEE Access 2021, 9, 68609–68618. [CrossRef]*
6. *Porwal, S.; Ostwal, G.; Phadtare, A.; Pandey, M.; Marathe, M.V. Sarcasm detection using recurrent neural network. In Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 14–15 June 2018; pp. 746–748.*
7. *Davidov, D.; Tsur, O.; Rappoport, A. Semi-supervised recognition of sarcasm in Twitter and Amazon. In Proceedings of the Fourteenth Conference on Computational Natural Language Learning, Uppsala, Sweden, 15–16 July 2010; pp. 107–116.*
8. *Bamman, D.; Smith, N. Contextualized sarcasm detection on Twitter. In Proceedings of the International AAAI Conference on Web and Social Media, Online, 8–10 June 2021; pp. 574–577.*
9. *Poria, S.; Cambria, E.; Hazarika, D.; Vij, P. A deeper look into sarcastic tweets using deep convolutional neural networks. arXiv 2016, arXiv:1610.08815.*
10. *Ilić, S.; Marrese-Taylor, E.; Balazs, J.A.; Matsuo, Y. Deep contextualized word representations for detecting sarcasm and irony. arXiv 2018, arXiv:1809.09795.*
11. *Ptáček, T.; Habernal, I.; Hong, J. Sarcasm detection on Czech and English Twitter. In Proceedings of the COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014; pp. 213–223.*
12. *Prasadu peddi, "The Adoption of a Big Data and Extensive Multi-Labled Gradient Boosting System for Student Activity Analysis" International Journal of All Research Education and Scientific Methods (IJARESM) ISSN: 2455-6211, Volume 3, Issue 7, July-2015, Impact Factor 2.287*
13. *Pawar, N.; Bhingarkar, S. Machine learning based sarcasm detection on Twitter data. In Proceedings of the 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 10–12 June 2020; pp. 957–961.*
14. *Bouazizi, M.; Ohtsuki, T.O. A pattern-based approach for sarcasm detection on Twitter. IEEE Access 2016, 4, 5477–5488. [CrossRef]*
15. *Tay, Y.; Tuan, L.A.; Hui, S.C.; Su, J. Reasoning with sarcasm by reading in-between. arXiv 2018, arXiv:1805.02856.*
16. *Joshi, A.; Bhattacharyya, P.; Carman, M.J. Sarcasm Detection Using Incongruity within Target Text. In Investigations in Computational Sarcasm; Springer: Berlin/Heidelberg, Germany, 2018; pp. 59–91.*
17. *Sriram, B.; Fuhry, D.; Demir, E.; Ferhatosmanoglu, H.; Demirbas, M. Short text classification in Twitter to improve information filtering. In Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information*

Retrieval, Geneva, Switzerland, 18–23 July 2010; pp. 841–842.

18. *Burfoot, C.; Baldwin, T. Automatic satire detection: Are you having a laugh? In Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, Suntec, Singapore, 4 August 2009; pp. 161–164.*
19. *Campbell, J.D.; Katz, A.N. Are there necessary conditions for inducing a sense of sarcastic irony? Discourse Process. 2012, 49, 459–480. [CrossRef]*
20. *Rajadesingan, A.; Zafarani, R.; Liu, H. Sarcasm detection on Twitter: A behavioral modeling approach. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 97–106.*
21. *Prasadu Peddi (2015) "EXPLORING THE IMPACT OF DATA MINING AND MACHINE LEARNING ON STUDENT PERFORMANCE", International Journal of Emerging Technologies and Innovative Research, ISSN:2349-5162, Vol.1, Issue 6, page no. pp314-318, November-2014, Available at : http://www.jetir.org/papers/JETIR1701B47.pdf*
22. *Shrikhande, P.; Setty, V.; Sahani, A. Sarcasm detection in newspaper headlines. In Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), Rupnagar, India, 26–28 November 2020; pp. 483–487.*
23. *Misra, R.; Arora, P. Sarcasm detection using hybrid neural network. arXiv 2019, arXiv:1908.07414.*
24. *Kumar, A.; Sangwan, S.R.; Singh, A.K.; Wadhwa, G. Hybrid deep learning model for sarcasm detection in Indian indigenous language using word-emoji embeddings. Trans. Asian Low-Resour. Lang. Inf. Process. 2022. [CrossRef]*
25. *Govindan, V.; Balakrishnan, V. A machine learning approach in analysing the effect of hyperboles using negative sentiment tweets for sarcasm detection. J. King Saud Univ.-Comput. Inf. Sci. 2022, 34, 5110–5120. [CrossRef]*
26. *Vinoth, D.; Prabhavathy, P. An intelligent machine learning-based sarcasm detection and classification model on social networks. J. Supercomput. 2022, 78, 10575–10594. [CrossRef]*
27. *Nayak, D.K.; Bolla, B.K. Efficient deep learning methods for sarcasm detection of news headlines. In Machine Learning and Autonomous Systems, Proceedings of ICMLAS 2021, Tamil Nadu, India, 24–25 September 2021; Springer: Singapore, 2022; pp. 371–382.*
28. *Guibon, G.; Ermakova, L.; Seffih, H.; Firsov, A.; Noe-Bienvenu, G.L. Multilingual fake news detection with satire. In Proceedings of the Computational Linguistics and Intelligent Text Processing: 20th International Conference, CICLing 2019, La Rochelle, France, 7–13 April 2019; Revised Selected Papers, Part II; pp. 392–402.*
29. *Tan, Y.Y.; Chow, C.-O.; Kanesan, J.; Chuah, J.H.; Lim, Y. Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning. Wirel. Pers. Commun. 2023, 129, 2213–2237. [CrossRef] [PubMed]*
30. *Alturayeif, N.; Luqman, H.; Ahmed, M. A systematic review of machine learning techniques for stance detection and its applications. Neural Comput. Appl. 2023, 35, 5113–5144. [CrossRef] [PubMed]*
31. *Goel, P.; Jain, R.; Nayyar, A.; Singhal, S.; Srivastava, M. Sarcasm detection using deep learning and ensemble learning. Multimed. Tools Appl. 2022, 81, 43229–43252. [CrossRef]*
32. *Misra, R. (Ed.) News Headlines Dataset For Sarcasm Detection; Kaggle: San Francisco, CA, USA, 2019. Available online: https://www.kaggle.com/datasets/rmisra/news-headlines-dataset-for-sarcasm-detection*
33. *Oraby, S.; Harrison, V.; Reed, L.; Hernandez, E.; Riloff, E.; Walker, M. Creating and characterizing a diverse corpus of sarcasm in dialogue. arXiv 2017, arXiv:1709.05404.*