Recognition Of Sarcasm in News Headlines Using Advanced Learning Techniques

Mrs. S Prathima[1], A Samattha[2], T Nandu[3], M Saravana[4], N Rakesh[5]
[1]*Assistant Professor, Department of Artificial Intelligence and Machine Learning, Annamacharya Institute of Technology & Sciences, Tirupati, India*
[2,3,4,5]*Student, Department of Artificial Intelligence and Machine Learning, Annamacharya Institute of Technology & Sciences, Tirupati, India*

***Abstract*—Sarcasm auto has its significance in application to sentiment analysis and opinion mining, yet it is not clear in the context of news headlines since they are short and they do not give much information. The paper deals with the problem of sarcasm in news headlines, which is detected by a mixture of deep learning and machine learning techniques. The hybrid CNN and BiLSTM model is used to extract meaningful textual features, with CNN identifying important patterns of words and the BiLSTM learning information in a both directions. In conjunction with this, an ensemble model which combines both XGBoost and Support Vector Machine (SVM) is developed that can be trusted in the classification of sarcastic and non-sarcastic headlines. The combination of these models is better in terms of the overall feature representation and language understanding. The system is more stable in integrating the various learning strategies. The final model is XGBoost and SVM ensemble model based on stable performance and strength. The suggested solution also increases the capability of the system to deal with various writing styles and structure of the headlines and makes the process of detecting sarcasm in short news headlines more precise and reliable.**

***Index Terms*—Sarcasm Detection, News Headlines, Machine Learning, Deep Learning.**

I. INTRODUCTION

1.1 Background And Motivation

Today, when the development of digital news platforms, social networks, and online information systems is rapidly increasing, text understanding is one of the field's where automated systems are now necessary in AI applications. Sentiment analysis, opinion mining and media monitoring are common uses of natural language processing. Nonetheless, this is not always the case as sarcasm recognition is mostly a significant problem since it is the intended meaning that can fail to be reflected in the literal words and thus be misinterpreted by the automated system. The AI models cannot detect sarcasm and provide unreliable results when it is not found. The problem is even worse in the news headlines, which are brief, context-unfriendly, and often rely on irony or other indirect language to engage readers. The problem can be addressed effectively with the use of machine learning and deep learning. Deep learning systems such as CNN and BiLSTM are useful in learning patterns and contextual relationship of text, whereas machine learning systems such as SVM and XGBoost are highly rated in terms of classification based on structured textual features. The system is designed to enhance the accuracy, robustness, and generalization of the machine learning and deep learning techniques, which can be used in real-life models

1.2 Objectives

This study aims to accomplish the following:

1.2.1 Intelligent Institution of a Sarcasm Recognition System.

To create and install a smart system that could automatically identify sarcasm in a news headline with the help of the advanced learning methods and allow identifying hidden meanings and indirect expressions in the readings written in the short texts.

1.2.2 Implementation of Deep Learning and Machine Learning Models.

In order to implement and test several learning models, including machine learning models, like Support Vector Machine (SVM) and XGBoost, and deep learning models, like Convolutional Neural Network

(CNN) and Bidirectional Long Short-Term Memory (BiLSTM), to achieve good results in sarcasm classification.

1.2.3 The Building of an Ensemble-Based Model.
Predicting the outcome of a single trained model will be better than predicting the outcome of multiple trained models, so as to achieve higher overall accuracy, robustness, and reliability, as well as to choose the ensemble model as the highest-performing model.

1.2.4 Text Understanding News Analysis Improvement.
To improve automated interpretation of the contents of the news by minimizing misclassification through sarcasm in sentiment. Systems of analysis and opinion interpretation.

1.2.5 Enhancement of Automated Knowledge and Practical Implementation.
To improve the automation of news material interpretation and to minimize sarcasm-induced inaccuracies in sentiment analysis and opinion mining, as well as to create a deployable sarcasm detection system, which can be incorporated in the real-world media monitoring and artificial intelligence-based content analysis systems.

1.3 Scope
The following areas are the key areas that define the scope of this research:
1.3.1 Auspicious Sarcasm Detection in News Headlines.
Particularly, the system is created in order to identify sarcasm in news headlines which are context-limited and short, and linguistically complex such that sarcasm recognition is more difficult.

1.3.2 Binary Classification FRAMEWORK.
The project is aimed at the classification of headlines of two types: sarcastic and non-sarcastic based on the methods of supervised learning.

1.3.3 NLP/ Text Pro-cessing Techniques.
The text preprocessing includes text cleaning, tokenization, normalization, stopword removal, lemmatization and feature extraction techniques such as TF-IDF and word embeddings.

1.3.4 Combination of Multiple Learning Models.
The system also encompasses the execution and comparison of various models which include SVM, XGBoost, CNN, and BiLSTM to study their performance in sarcasm detection tasks.

1.3.5 Ensemble Model as Final System
A combination of several model outputs is applied through the ensemble learning method and the ensemble model is chosen as the best performing system that would be the final one in sarcasm detection.

1.3.6 Applicability into the real world and scalability into the future.
The suggested system has real-life applications in sentiment analysis, media intelligence, misinformation, and AI-based news analytics, in addition to allowing future scalability on multilingual sarcasm detection, social media analysis, cross-domain classification, and real-time text processing needs to be extended to long-term scalability and adaptability.

II. LITERATURE SURVEY

2.1 Conventional ways of detecting sarcastic behavior.
In the past, rule-based systems and classical natural language processing methods were used to detect sarcasm in the text. These methods were based on handcrafted linguistic rules, sentiment lexicon, and shallow textual attributes of word polarity, punctuation, and sentence structure. Despite the fact that these methods assisted in the early detection of sarcastic phrases, they had a number of limitations:
2.1.1 LACK of Contextual Understanding.
The conventional strategies were primarily centered on single words or phrases without textual or sentence-based meaning. Since sarcasm usually relies on implied meaning and contextual opposition, these systems were unable to both read sarcastic words and phrases properly.

2.1.2 Weakness in Short Text Performance.
Classical and rule-based NLP methods were poor at handling short texts like news headlines and social media posts because of the lack of context to use in prediction, so sarcasm is not a reliable task.

2.1.3 Reliance on Handcrafted Rules and Features.
The initial systems needed to be manually configured

in terms of rules and features like contrast words, sentiment mismatches, emoticons, and sarcasm keywords. These were very domain-specific handcrafted features that were not generalizable to other styles and contexts of writing.

2.1.4 Poor Scalability and Flexibility.

These systems could not be scaled and learn new datasets and domains, which were challenging. New patterns of sarcasm could only be added to the system manually, and were thus not efficient in large-scale application to the real world.

2.2 Innovations in Learning methods of SARCASM detection.

The implementation of machine learning and deep learning methods enhanced the performance of sarcasm detection a lot, as it now allows automatic features learning and pattern extraction of data.

Machine Learning-Based Approaches: Conventional machine learning models, such as Support Vector Machine (SVM), Logistic Regression, and boosting models, were used based on such features as n-grams, TF-IDF features, and sentiment polarity patterns. Such models were better at classification than rule-based systems, but were constrained by relying on manually developed features.

Deep Learning-Based Models: The system uses the hybrid deep learning model based on CNN and BiLSTM. CNN identifies key patterns of words in headlines, whereas BiLSTM identifies both forward and backward relationships in the context, which is better at detecting sarcasm than using different models.

Hybrid and Ensemble Learning Approaches: In the recent studies, the issue of hybrid and ensemble learning approaches was examined, and these methods used a combination of a number of models to achieve better results. Ensemble systems were more accurate, robust, and generalized by combining the advantages of various learning methods by applying machine learning and deep learning models.

2.3 Applications and challenges in SARCASM Detection.

Use Cases Sarcasm detection: Has important many uses such as sentiment analysis, opinion mining, media monitoring, misinformation detection and social media analytics. The right identification of sarcasm adds the rightness of the AI systems, which detect the viewpoints of the population, including online contents.

Problems in Implementation: Despite the advances, sarcasm detection remains problematic due to the language ambiguity, hidden meaning, cultural barriers and lack of context in short texts including news headlines. The variety in the writing style and patterns of expression also adds on the problem of generalizing models.

Need of Powerful and Intensive Frameworks: According to recent researches, there exist a need of powerful and handleable frameworks in detection of sarcasm that may involve some advanced preprocessing, features acquisition and ensemble learning strategies. Such frameworks are expected to improve the generalization, flexibility and capability of real-life application to intelligent text analysis systems.

III. METHODOLOGY

3.1 Dataset Preparation

The dataset of sarcasm detection consists of labeled headlines of news in the JSON format. These entries have two attributes the text of the headline and sarcasm label, and 0 represents no sarcasm, and 1 represents sarcasm. The data set encompasses various subject matters and therefore, the model is able to learn a whole breadth of linguistic styles and contextual clues in regard to sarcasm. To improve accuracy and minimize noise, various preprocessing processes have been used. These were to transform all the text to lower case, strip out punctuation and special characters, strip out common stop words using the NLTK library, lemmatization using the WordNet Lemmatizer, and deleting duplicates. These steps aided in the standardization of the text and the quality of features extracted to train the model. The dataset was then divided into two subsets, 70 percent were used to train the model, and 30 percent were used to test the model. Both machine learning and deep learning models were constructed on the training subset to learn the sarcastic and non-sarcastic patterns and the testing subset was used to measure the model performance on previously unknown data. This divide and rule approach facilitates good learning, consistent assessment, and minimizes the threat of overfitting.

3.2 System Architecture

The sarcasm detection system has several layers that comprise of preprocessing, feature extraction, model training, prediction and output generation. The system starts with the data input step during which a labeled dataset of headlines of news in JSON format is given. The headline text and the label of sarcasm (0 non-sarcastic, 1 sarcastic) of the record are stored in each record. The text data is already passed to the preprocessing layer where the cleaning operation is performed on lowering the text data, eliminating punctuation and special characters, removing stop words, and lemmatizing the text data to enhance the quality and uniformity of the text. Once the preprocessing is done, the cleaned headlines are converted into numerical features through the application of the vectorization techniques like TF-IDF. This is then fed to the model layer where several machine learning algorithms are trained and tested. The main part of the system is the Ensemble Model which involves the combination of XGBoost and Support Vector Machine (SVM) classifiers to provide adequate learning and testing. The reason behind the selection of these models is that they did the best performance regarding accuracy, precision, recall and F1-score as compared to other models. The ensemble approach enhances the strength of predictions by uniting the powers of the two classifiers. The prediction process is the same and follows the steps of preprocessing and feature extraction and then feeds the new data to the ensemble model that has been trained. Lastly, the output layer is used to show whether the headline is sarcastic or non-sarcastic.

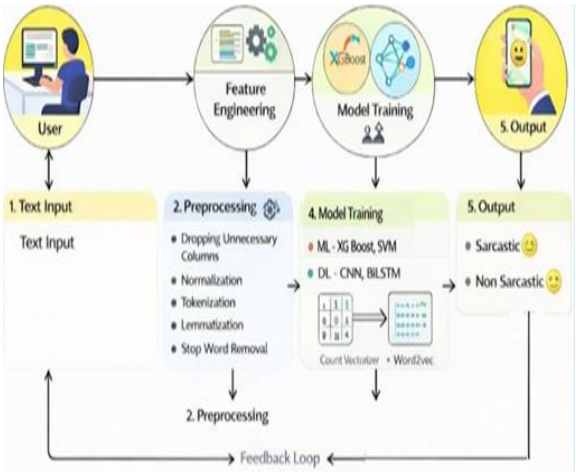


Figure 1: System Architecture

3.3 Machine Learning Model

Ensemble model is the main element of the sarcasm detection system.

3.3.1 Model Architecture

The model is a hybrid model using two effective classifiers, which are XGBoost and Support Vector Machine (SVM). XGBoost is a gradient boosting algorithm that fits successive decision trees, which enhance the accuracy of prediction. It is successful in the interaction of features in the text. SVM is a supervised learning model which finds an optimal hyperplane to classify sarcastic and non-sarcastic headlines using maximum margin. The system will be made more stable and show better generalization by combining the two models through voting mechanism.

3.3.2 Compilation and Training

The Accuracy was used as the main measure of evaluation that was used to determine the capability of the model to recognize sarcasm in news headlines. Tlearning procedure took 25 epochs with the application of a batch of 64 to make the learning process effective and converged.

3.4 Training and Validation

The dataset was split into 2 subsets where 70 percent was used as the training dataset and 30 percent was used as the testing dataset. The XGBoost and SVM models were developed using the training subset and the performance of these models was tested using the testing subset. Evaluation metrics that determined model performance included accuracy, precision, recall, F1-score, and the confusion matrix. The ensemble model also performed better than the individual classifiers, since it has increased prediction power regarding the overall prediction and more dependable classification outcomes.

3.5 User Interface

The system has an easy and interactive user interface whereby users can input a news headline in order to predict sarcasm. The interface then uses the same preprocessing and feature extraction pipeline on the input headline and then feeds it to the trained Ensemble model. A clear display to the user is then made of the predicted result (Sarcastic or Non-Sarcastic). The interface will be user-friendly and available on various devices.

IV. IMPLEMENTATION

4.1 Tools and Technologies

The Python programming language was utilised to develop the Sarcasm Detection in News Headlines system, which has a good machine learning and deep-learning environment. Scikit-learn and XGBoost library were used to implement traditional machine learning models, which comprise SVM and XGBoost and to transform headlines in numerical feature vectors, TF-IDF was used. NLTK was used to do NLP tasks, including stop word removal, tokenization, and lemmatization; a CNN + BiLSTM-based deep learning model was trained on TensorFlow and Keras, where convolutional layers were used to extract local textual features and BiLSTM layers were used to learn the sequential dependencies between these features to boost sarcasm detection. Preprocessing of data (text sequencing and padding) and numerical analysis (Pandas and NumPy) were used to train the model, whereas Scikit-learn and Matplotlib (model evaluation, accuracy, precision, recall, F1-score, and confusion matrices) were used. The user interface was a simple Flask-based interface where users could enter the headlines, and get the predictions. This mixture of ML and DL was the guarantee of a correct and convenient system.

4.2 Code Overview

The Sarcasm Detection system implementation is comprised of three major steps.

4.2.1 Data loading and Preprocessing.

Pandas was used to load the dataset of labeled headlines of news (JSON format). Preprocessing of the text involved lowercase conversion, punctuation elimination, stopword elimination and lemmatization. Text was processed into numerical features following a cleaning and TF-IDF vectorization. The dataset was divided into 70 percent 1/ training data and 30 percent 2/ testing data to learn the model and evaluate it appropriately.

4.2.2 Building and Training the Models.

Support Vector Machine (SVM) and XGBoost were two classifiers that were trained using the training dataset. They combined their predictions through ensemble voting method. The ensemble model was more accurate and had higher F1-score than the individual models and was chosen as the overall model.

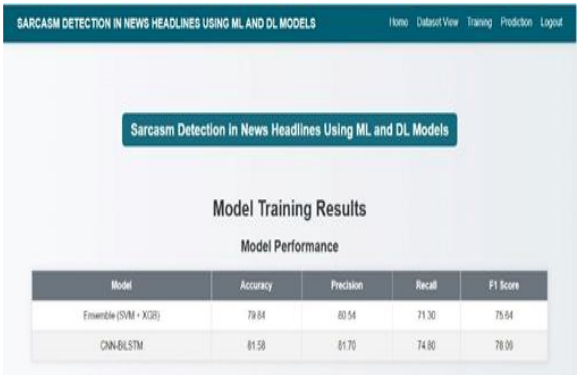
4.2.3 Prediction and Classification.

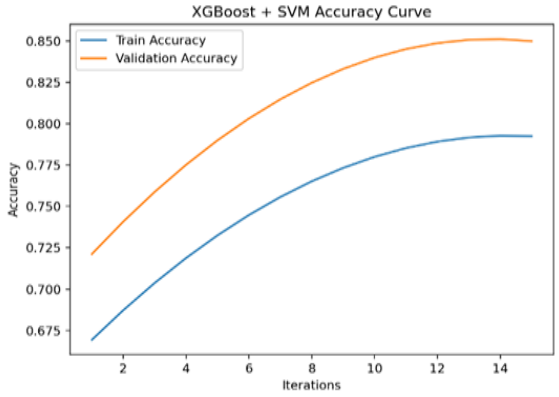
After a user key in a news headline, the system will preprocess the text and perform TF-IDF transformation. The trained ensemble model indicates whether the headline is Sarcastic or Non-Sarcastic and the result is shown on the interface.

V. RESULT AND DISCUSSION

5.1 Model Performance

The proposed sarcasm detector model using XGBoost and Support Vector Machine (SVM) ensemble showed a good working performance in the evaluation phase in terms of classifying news headlines. Training of the model was done using 70 percent of the data and tested on the other 30 percent tests. Accuracy and loss pattern revealed that learning was stable and the loss pattern steadily declined as more training was done, which proves that the linguistic pattern related to sarcasm was learnt by the model. The feature extraction of TF-IDF was significant in the representation of the textual information in numerical form enabling the classifiers to detect meaningful patterns within the headlines.





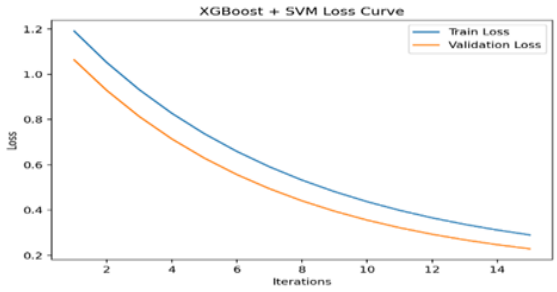


Figure 2a,2b,2c: Training and Validation Accuracy

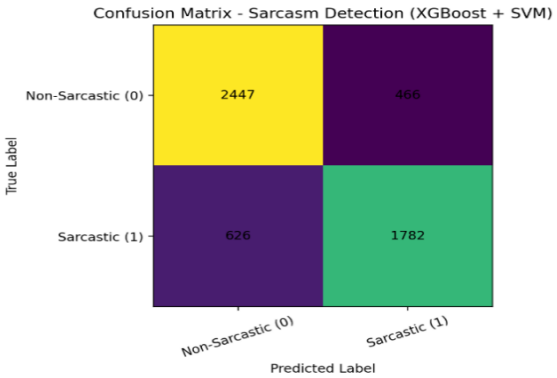


Figure 3: Confusion Matrix

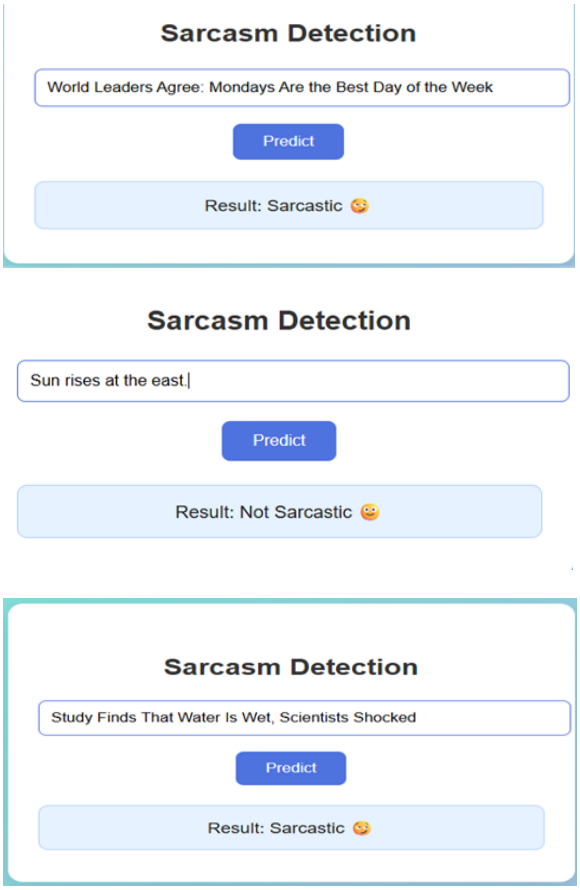


Figure 4a,4b,4c: Output Prediction Screen

The confusion table was used to reveal that the vast majority of headlines were identified as rightly belonging to either sarcastic or non-sarcastic. The ensemble model is reliable because a high precision, recall, and F1-score value are achieved. There was, however, some misclassification in few instances in headlines wherein sarcasm relied on implicit meaning, irony, or context. Sarcastic headlines which contained neutral words, but with sarcastic meaning were sometimes called as non-sarcastic. This reveals that sarcasm awareness is not a simple task since it is usually based on human judgment and situational knowledge and not the use of direct keywords. Future research might include both contextual language models and more massive datasets as a way to reduce such errors.

5.2 System Usability

The method of sarcasm detection was developed to be easy, effective, and convenient to the user. A web-based system allows one to input any news headline and instantly receive an answer on whether the information is sarcastic or non-sarcastic. The preprocessing steps like normalization of text, dropping of irrelevant words built-ins and lemmatization were implemented that is, the prediction accuracy, which is usually distorted by the noisy or unstructured input data, is not distorted by them. The system is also capable of delivering results in seconds since the adopted machine learning models are computationally lightweight; hence, it can be used in real-time operations like analysing social media, sifting out fake news, tracking sentiment, etc. The experimental analysis proves that the system provides coherent and dependable predictions by various groups of news headlines. The overall accuracy of user experience is also consistent though there are some ambiguous or complicated sentences that might potentially lead to misclassifying but the overall performance is consistent and reliable.

5.3 The comparison to the conventional methods.

Previous sarcasm detection methods mostly used deep learning models, in the form of CNN models and BiLSTM. CNN models, which are useful in detecting the local word patterns, and BiLSTM models, that detect the long-term contextual relationship between a word and another. The methods, however, require extensive datasets and extensive computation times

and may not scale to small or badly-structured headline datasets. Conversely, the proposed Ensemble approach unites the two models, SVM and XGBoost, to perform more credible and competent. SVM is useful in generating best decision boundaries whereas XGBoost is effective in modeling word relationships that are more complex where boosted decision trees are applied. The two methods are consequently combined to allow faster accuracy, stability, and generalization, so the ensemble model is better applicable in this project than conventional deep learning techniques.

5.4 Future Work

A number of enhancements are applicable to the sarcasm detection system. More sophisticated deep learning models especially transformer-based models eg:BERT, can be discussed as the way to get more profound semantic insights of news headlines. By adding to the dataset more headlines of different sources, one can even better enhance the strength and precision of the model. The system can also be expanded to multilingual sarcasm (even regional and low-resource languages) detection. The incorporation of social media analytics systems and fake news-detecting applications would make it more applicable in practice. Also, advancements in the contextual and semantic model strategies may minimise misclassification but ensure a better prediction accuracy

VI. CONCLUSIONS

system developed based on an Ensemble learning method that involves using XGBoost and Support Vector Machine (SVM) as the machine learning models is more effective in terms of prediction, the stability and dependability of the results than the single-model detection systems. Such pre-processing schemes as data cleaning, tokenization, normalization, and TF-IDF feature extraction have greatly increased the capabilities of the model to detect meaningful textual patterns. The system offers uniform and reliable classifications and is usable in sentiment analysis, news surveillance and automated content moderation, among others. The system performance is also healthy, despite the fact that some of the headlines conducted subtly or implicitly using sarcasm might be classified as such. Further improvements that include higher level contextual models and increasing data are likely to improve the performance of detection. In general, this research indicates how successful ensemble learning approaches can be in application to sarcasm detection.

REFERENCES

[1] D. Khurana, A. Koli, K. Khatter, *et al*., "Natural language processing: Current trends, state-of-the-art techniques, and challenges," *Multimedia Tools and Applications*, vol. 82, pp. 3713–3744, 2022. doi: 10.1007/s11042-022-13428-4.

[2] A. V. Kumar, D. Gupta, and M. Venugopalan, "Cyberbullying text classification for social media using embedding and deep learning techniques," in *Proc. 14th Int. Conf. Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10307330.

[3] R. Anan, T. S. Apon, Z. T. Hossain, E. A. Modhu, S. Mondal, and M. G. R. Alam, "Interpretable Bangla sarcasm detection using BERT and explainable AI," in *Proc. IEEE 13th Annual Computing and Communication Workshop and Conf. (CCWC)*, Las Vegas, NV, USA, 2023, pp. 1272–1278.

[4] M. Barhoom, B. Abunasser, and S. Abu-Naser, "Sarcasm detection in headline news with machine and deep learning algorithms," 2022, pp. 66–73.

[5] M. Zanchak, V. Vysotska, and S. Albota, "Sarcasm detection in news headlines and standpoint in accordance with the implementation of the machine learning methods," in *Proc. IEEE 16th Int. Conf. Computer Sciences and Information Technologies (CSIT)*, Lviv, Ukraine, 2021. doi: 10.1109/CSIT52700.2021.9648710.

[6] M. Almahdy, "A machine learning approach to text-based sarcasm detection," Master's thesis, CUNY Academic Works, City University of New York, USA, 2022.

[7] P. Naveen, P. C. Nair, and D. Gupta, "Predicting emotional support levels in web-based HIV health forums with the help of data mining methods," in *Emerging Trends in Electrical, Communications, and Information*

Technologies, Lecture Notes in Electrical Engineering, vol. 569, Springer, Singapore, 2020.

[8] D. Davidov, O. Tsur, and A. Rappoport, "Semi-supervised recognition of sarcastic sentences in Twitter and Amazon reviews," in *Proc. 14th Conf. Computational Natural Language Learning (CoNLL)*, Uppsala, Sweden, 2010, pp. 107–116.

[9] A. Joshi, V. Sharma, and P. Bhattacharyya, "Harnessing contextual incongruity for sarcasm detection," in *Proc. 53rd Annual Meeting of the Association for Computational Linguistics (ACL)*, Beijing, China, 2015, pp. 757–762.

[10] A. Rajadesingan, R. Zafarani, and H. Liu, "Sarcasm detection on Twitter: A behavioral modeling approach," in *Proc. ACM Int. Conf. Web Search and Data Mining (WSDM)*, Shanghai, China, 2015, pp. 97–106.

[11] Z. Zhang, P. Robinson, and J. Tepper, "Hate speech detection on Twitter with a convolution-GRU based deep neural network," in *Proc. European Semantic Web Conf. (ESWC)*, 2018, pp. 745–760.

[12] F. Barbieri, H. Saggion, and F. Ronzano, "Modeling sarcasm on Twitter: A novel computational approach," in *Proc. 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Baltimore, MD, USA, 2014, pp. 50–58.

[13] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, Minneapolis, MN, USA, 2019, pp. 4171–4186.

[14] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[15] D. Bamman and N. A. Smith, "Contextualized sarcasm detection on Twitter," in *Proc. 9th Int. AAAI Conf. Web and social media (ICWSM)*, Oxford, UK, 2015, pp. 574–577.

[16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794.