

The Achilles 'heel of Sentiment Analysis: Sarcasm Detection

Dr. R.K. Dhuware¹, Ms. Pratibha Jaisingh²

¹. Head , Department of Computer Science, Dhote Bandhu Science College, Gondia

². Research Scholar, Department of Electronics and Computer Science, RTM Nagpur University, Nagpur.

Abstract: Sarcasm is the cruel application of irony, in which words are used to convey the antithesis of their apparent meaning in a playful or mocking manner. When someone uses sarcasm, they are expressing their feelings or opinions in an oblique manner rather than outright. The objective of sarcasm detection is to identify sarcasm. Sarcasm detection is incredibly tough for machines, since it is quite difficult for humans to comprehend sarcasm in people around them. Sarcasm detection is one of the issues that is currently popular in order to stop misconceptions in communication, especially online, and to progress and improve sentiment analysis. In this paper, we discuss several types of sarcasm, the process of sarcasm detection, different methods of sarcasm detection, problems, difficulties, and the potential for sarcasm detection in the future.

Keywords: Sarcasm Detection, Sentiment Analysis, Machine Learning,

I. Introduction

Internet and social media use have surged since the Corona virus hit the planet. People are allowed to share their opinions about anything, including their satisfaction with a product, their criticism of a movie, their rage at a criminal act, their disgust with government services, etc. Many companies place a high value on public opinion, particularly commercial firms that regularly seek client feedback to assess the standing of their goods. Sentiment analysis plays a very significant role for such firms because Sentiment analysis is a process of analysing whether a text contains positive, negative or neutral review about certain product or services. However, some individuals comment in opposite to what they really want to express. That is, they make use of sarcasm to insult or mock the organizations. . Sarcasm is kind of sentiment where a person will not tell directly about their opinion or feeling, but they express it indirectly. Therefore, it is difficult to use sentiment analysis techniques to detect sarcasm. Hence, sarcasm detection is called Achilles' heel of sentiment analysis.

As defined in oxford dictionary, Sarcasm is a way of using words that are the opposite of what you mean in order to be unpleasant to somebody or to make fun of them ^[1]. Cambridge Dictionary defines Sarcasm as the use of remarks that clearly mean the opposite of what they say, made in order to hurt someone's feelings or to criticize something in a humorous way^[2] The Collins Dictionary defined Sarcasm as “**Sarcasm** is speech or writing which actually means the opposite of what it seems to say. Sarcasm is usually intended to mock or insult someone” ^[3]. In verbal communication, where we may infer someone's intentions from their tone of voice or facial expression, sarcasm may be simple to spot. Sarcasm may be simple for a human to pick up on in text communication. However, since positive terms are frequently included in sarcastic statements, it will be challenging for machines to recognise sarcasm automatically. For instance, "I feel out of this world to do office job on Sunday" uses a good mood yet has unfavourable intentions. Human intelligence can grasp this sarcasm, but machines will be confused.

In this paper, we examine an empirical investigation of sarcasm detection. The remainder of the paper is structured as follows: Types of sarcasm are explored in section II. Section III discusses the steps for sarcasm detection. Feature Extraction and feature selection are covered in Section IV. Section V discusses various sarcasm detection methods. In section VI, Major issues and difficulties are covered. Finally, we wrap up with a few thoughts and plans for further sarcasm detection research in sect VII.

II. Types of sarcasm

Numerous researchers categorise various forms of sarcasm based on the features and textual structures that best suit their point of view. We summarized the sarcasm into three types according to Rajadesingan et al. [4], Bharti et al. [5] and Camp [6] given in Table I.

T1	<ul style="list-style-type: none"> a) Sarcasm as a contrast of sentiments b) Sarcasm as a complex form of expression c) Sarcasm as a means of conveying emotion d) Sarcasm as a possible function of familiarity e) Sarcasm as a form of written expression
T2	<ul style="list-style-type: none"> a) Contrast between positive sentiment and negative situation b) Contrast between negative sentiment and positive situation c) Fact Negation - text contradicting a fact d) Likes and Dislikes Prediction – behaviour based e) Lexical Analysis - sarcasm hashtag based f) Temporal Knowledge Extraction - tweets contradicting facts about event
T3	<ul style="list-style-type: none"> a) Sarcasm and Verbal Irony b) Propositional Sarcasm c) Lexical Sarcasm d) Like'-prefixed Sarcasm e) Illocutionary Sarcasm

The types of Sarcasm given in table I are explained below
T1.

a) Sarcasm as a contrast of sentiments:

In this type of sarcasm, the sentence contains positive emotions like "love," "enjoy," and "etc.," it depicts an unpleasant or negative experience. It is most commonly used form of sarcasm. For example, in sentence “wow, it’s Sunday morning and I am doing office work”, positive sentiment wow is used followed by sad state of working on Sunday morning.

b) Sarcasm as a complex form of expression

The ability to express sarcasm and cognitive complexity are positively correlated in a small but significant way [4]. Making sound decisions requires a high level of cognitive complexity, which includes comprehending and taking into account various points of view. Additionally, in order to convey sarcasm, one must ascertain whether the situation is right for it, come up with a proper sarcastic phrase, and determine whether the recipient is likely to understand it. So, sarcasm is a complicated style of expression that requires more effort from the user than usual.

c) Sarcasm as a means of conveying emotion

In this type of sarcasm, user expresses his/her emotions according to their mood in sarcastic way. Exceedingly happy words are used by the user when they are extremely irritated or angry. An illustration, “What an AMAZING day! I love working overtime”. Person is expressing extreme happiness because of overtime but actually person is annoyed.

d) Sarcasm as a possible function of familiarity

Sarcasm is utilised in this form if the speaker is familiar with the surroundings or the people they are living with. For example, (Wife to Husband after husband forgot to take out the trash) “Gosh! I love when our house looks and smells so clean.” Here wife is familiar with husband’s nature of forgetting things.

e) Sarcasm as a form of written expression

Sarcasm in writing has become more prevalent with the rise of social media platforms. In order to convey the opposite idea, the user displays their negative sentiment in writing by utilising repeated characters or punctuation. For instance, “Wowwww!!!! The food is mildly spicy here”. Here, the user is trying to say that the food is quite spicy. Including the hashtag #sarcasm in the text is another form of textual sarcasm. For example, “I love being alone” #sarcasm. It’s mostly used in social media.

T2

a) Sarcasm as contrast between positive sentiment and negative situation

This sarcasm arises from the juxtaposition of positive sentiment word with a negative activity or state [7]. Example, “Absolutely adore it when my bus is late”

b) Sarcasm as contrast between negative sentiment and positive situation

Here, sarcastic sentences includes negative sentiments followed by positive state or activity. For example, “I hate India in cricket, because they always win”. Here hate is negative word but the situation is positive.

c) Sarcasm as fact Negation

Here, sarcasm is utilised to express negative feelings while using pleasant language and to challenge an established truth or fact. “I enjoy waking up as the sun rises in the west”, for instance. Here person is actually annoyed because he is waking in the evening when sun sets in the west.

d) Likes and Dislikes Prediction – behaviour based

Here the user’s behaviour is expressed in terms of likes or dislikes about various factors like products, services, events etc. which indicated sarcasm.

e) Lexical Analysis - sarcasm hashtag based

This sarcasm is based on hashtag (#). The text contain #sarcasm in the end of the text.

f) Temporal Knowledge Extraction - tweets contradicting facts about event

This sarcasm is when a text conflict with the fact about any event such as a festival, birthday, sports, etc.

T3

a) Sarcasm and Verbal Irony

Verbal irony and sarcasm are often used interchangeably, but there is actually one key difference between the two. Sarcasm is always intentional when someone is being

sarcastic, they are deliberately poking fun at someone or something. In contrast, irony can be either intentional or accidental.

b) Propositional Sarcasm

Propositional sarcasm have implicit sentiments hidden inside a statement

To comprehend a sentence's sentiment, it is crucial to understand the context of the sentence. Saying "Your proposal sounds amazing," for instance, must actually be implying that the plan is horrible.

c) Lexical Sarcasm

Expressions like "brilliant," "inspired," "genius," "diplomat," and "thrilled" are frequently the targets of lexical sarcasm because they signify the extreme end of a commonly associated, normatively loaded scale. The sarcastic inversion then supplies a value at the scale's extreme other end. Often, the targeted expression denotes a positive value, but it can also be negative, as in "If you manage to generate one more half-baked, inconsequential idea like that, then you'll get tenure for sure"

d) Like'-prefixed Sarcasm

'Like'-prefixed sarcasm only combines with declarative sentences, and only targets content that is determined by composition of the constituent expressions' conventional meanings plus lexically-focused pragmatic processes.

e) Illocutionary Sarcasm

Illocutionary Sarcasm occurs when the sarcasm's scope embraces the complete illocutionary act that a sincere utterance of the relevant sentence would have undertaken, not just a particular element inside the spoken sentence or a premise related with the utterance. "How old did you claim you were, for instance?"

III. Steps of Sarcasm Detection

The main steps in Sarcasm Detection are:

- a) Data Collection
- b) Data Pre-processing
- c) Feature extraction
- d) Feature Selection
- e) Applying Sarcasm Detection Algorithms
- f) Evaluating Results

Figure (1) shows the steps of Sarcasm Detection

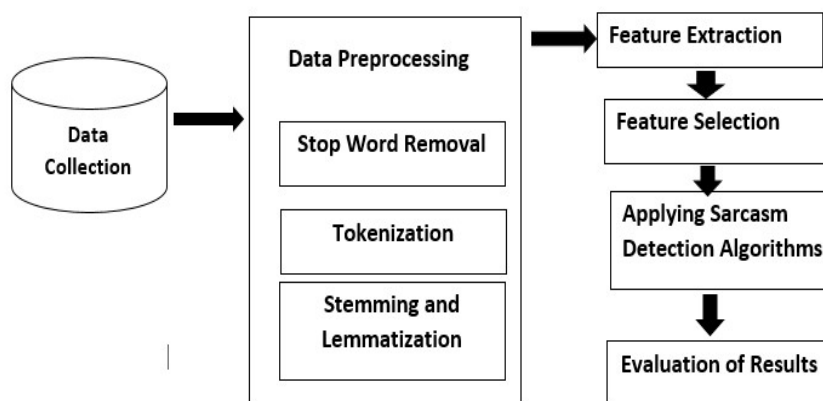


Fig (1) Steps in Sarcasm Detection

Each of the steps are discussed below

- 1) **Data Collection :**
Data acquisition is the first stage in any data modelling process. There are two methods for gathering information for sarcasm detection.
 - a) Application Programming Interface, or API, is a tool that enables user interaction with applications. Users can connect with Twitter, for instance, by utilising the Twitter API. Through the Twitter API, users can download tweets using hashtags. Twitter's #sarcasm hashtag can be used to identify sarcasm.
 - b) By utilising datasets like the Twitter dataset, SARC (Self-annotated Reddit Corpus), and news headline dataset, among others.
- 2) **Data Pre-processing:** Data should be pre-processed to eliminate noise after collection before being used as an input for a model. Data pre-processing involves the following steps:
 - a) **Stop words removal:** Stop words are common words in the English language such "as," "is," "and," "the," etc. These words have been deleted since they are not necessary for sarcasm detection.
 - b) **Word tokenization:** Tokenization refers to the breakdown of a sentence into its component words or tokens. Tokenization, for instance, will break a phrase like "I love music" into three words, such as "I", "Love", "Music."
 - c) **Stemming and lemmatization:** These two processes both produce the inflected word's root form. Stemming will change the three words "goes," "going," and "gone" into the word "go," but it will change the words "finally," "finalised," and "finally" into the word "fina." In order to make these three words form the word "final," lemmatization is used.

3) Feature Extraction:

In this step, feature are extracted and based on these features, forms a data representation which is appropriate for the NLP task or the classification model.

4) Feature Selection: The most appropriate features are selected in feature selection step. Step 4 and step 5 are discussed in details in section IV

5) Applying Sarcasm Detection Algorithm:

There are four different approaches for sarcasm detection:

- 1) Rule –based approach
- 2) Lexicon-based approach
- 3) Machine learning based approach
- 4) Deep learning based approach

Each of the approach is discussed in section V.

6) Evaluation of Result:

We will use precision, Recall and F-score to evaluate the performance of sarcasm detection model.

Precision is defined as proportion of rightly foretold sarcastic statements with the entire forecasted sarcastic statements.

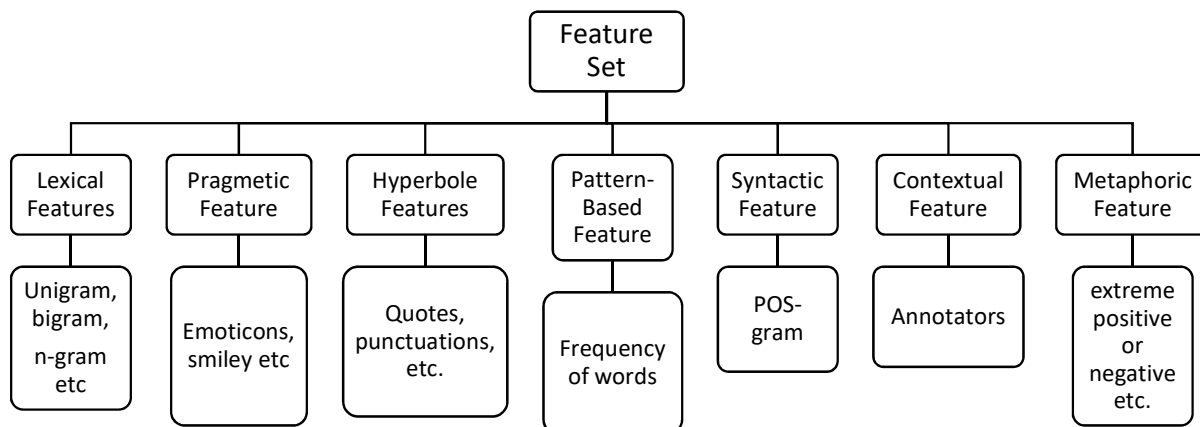
Recall is defines as the proportion of accurately forecasted sarcastic statements to all the genuine sarcastic statements present in the dataset.

F-score or f-measure is calculated using the values of Recall and precision as follows:

$$F\text{-score} = \frac{(2 * \text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precis})}$$

IV. Feature Extraction and Feature Selection in Sarcasm Detection Process

One of the most essential attributes of any NLP activity, including sentiment analysis and sarcasm detection, is features ^[8]. To forecast the more accurate output, various classification algorithms must be trained on well-defined features. We summarize the feature set in figure (2).



i) Lexical Feature:

Lexical feature include text properties like N-gram, Uni-gram and bi-gram. N-gram is a combination of N words. N means no. of words. For eg, 'sarcasm detection' is a 2-gram or bi-gram because it is combination of 2- words.

For the sentence "I love research ", In Unigram representation, the sentence will split into "I","Love","research". In bi-gram representation, two adjoining words will be considered together. "I love", "love research"

ii) Pragmatic Features:

Symbolic and figurative texts such as emoticons, smileys are example of pragmatic feature. Many researchers have used this features for sarcasm detection as presence of emoji's, smileys shows strong indication of sarcasm.

iii) Hyperbole Features:

Hyperbole is a figure of speech that creates heightened effect through deliberate exaggeration. Text properties like intensifiers, interjections, quotes, punctuation, etc. is called hyperbole. Presence of hyperbole is also a strong indication of sarcasm in the text. For e.g., "Great!!! I failed again in interview".

iv) Pattern-based Feature:

Some words appear more in the text than some other words. These high-frequency words can be used in sarcasm detection.

v) Syntactic Feature:

This feature contain recurrent sequence of morphosyntactic pattern (combination of morphology and syntax). Morphology is the study of words, how they are formed, and their relationship to other words in the same language. Example of syntactic feature is PoS-gram.

vi) Contextual Feature :

Contextual features are those features designed to identify sarcasm that draw on data outside of the text. In general, context refers to any knowledge or background information outside of the text to be predicted. It can be integrated utilising auxiliary data or details from the data source platform. Consider annotations. Additional context is requested by annotators in the form of reddit comments with sarcasm labels.

vii) Metaphoric Feature:

Extreme positive or negative nouns, extreme adjectives, proverbs, honorifics, etc. are all included in these features. This feature includes significant and useful data on the author's feelings.

Feature Extraction: If we have textual data, we cannot feed that data to any sarcasm detection algorithm since text data is not understood by the algorithm.

It exclusively comprehends numerical data. Feature Extraction from the text is the method for turning text data into number in the form of vector. Another name for it is text vectorization.

Techniques of Feature Extraction:

1) Bag-of-Words:

Bag-of-Words is simplest way of representing text. The data gathered is considered as a document in this instance, which can be anything from a sentence to a book page to a book chapter to a whole book. A single tweet or comment on a social media post could be the document for sarcasm detection. The input for Bag of Words is training data that includes a list of words that were discovered during data pre-processing. It counts how many times each word appears in the document, and the final word vector incorporates those counts for each word.

2) Term Frequency and Inverse Term Frequency (TF and IDF)

TF-IDF stands for term frequency-inverse-document frequency and gives more weight to words that are unique to a document than to words that are frequent but repeated throughout most documents. This allows us to give more weight to words uniquely characteristic to particular document.

Term Frequency (TF):

The number of times a word appears in a document is divided by the total number of words in that document. $0 < TF < 1$

Inverse Document Frequency (IDF):

The logarithm of the number of documents in the corpus is divided by the number of documents where the specific term appears

$$TF-IDF = TF * IDF$$

3) Word embedding like Word2vec and Glove

Representing word using word embedding is powerful because they are a result of training a neural network that predicts a word from all other words in the sentence.

Word2vec:

Word2vec is a technique for natural language processing (NLP) published in 2013. The word2vec algorithm uses a neural network model to learn word associations from a large corpus of text. It is an extremely efficient word embedding method for learning from unprocessed text input. Similar words in word2vec representation were mapped to nearby values. The relationship between various words and semantic information is upheld in word2vec word embedding. Word2vec, for instance, shows the words queen and ladies as related words.

Glove:

Another well-known word embedding technique is Glove (Global Vectors for Word Representations). Based on a method called matrix factorization, Glove creates a matrix

Feature Selection:

To improve the result of sarcasm detection, most important features are selected using feature selection methods. Popular methods for feature selection are chi-square and Mutual information.

Chi-square (χ^2):

Categorical features in a dataset are tested using the chi-square method. We determine the Chi-square between each feature and the target feature (Sarcasm or not) and then choose the features with the highest Chi-square scores. Because a greater chi-square value denotes a stronger relationship between a feature and expected value. The chi-square approach use a contingency table to examine the values. Chi-square testing is mostly used to determine if two category variables are statistically dependent or independent. Creating a hypothesis is the first step in the chi-square test. Following the definition of the hypothesis, the contingency table is created. It is important to determine the expected values before calculating the chi-square values. The null hypothesis is either accepted or rejected based on chi-square values.

Mutual Information:

Mutual information (MI) is a non-negative value that expresses how dependent two random variables are on one another. Higher values indicate greater dependence, and it equals 0 only when two random variables are independent.

V. Sarcasm Detection Techniques:

There are four different approaches for sarcasm detection:

- 1) Rule –based approach
- 2) Lexicon-based approach
- 3) Machine learning based approach
- 4) Deep learning based approach

AbouBaker ^[8] has suggested 4 methods to detect sarcasm which are shown in the fig (3).

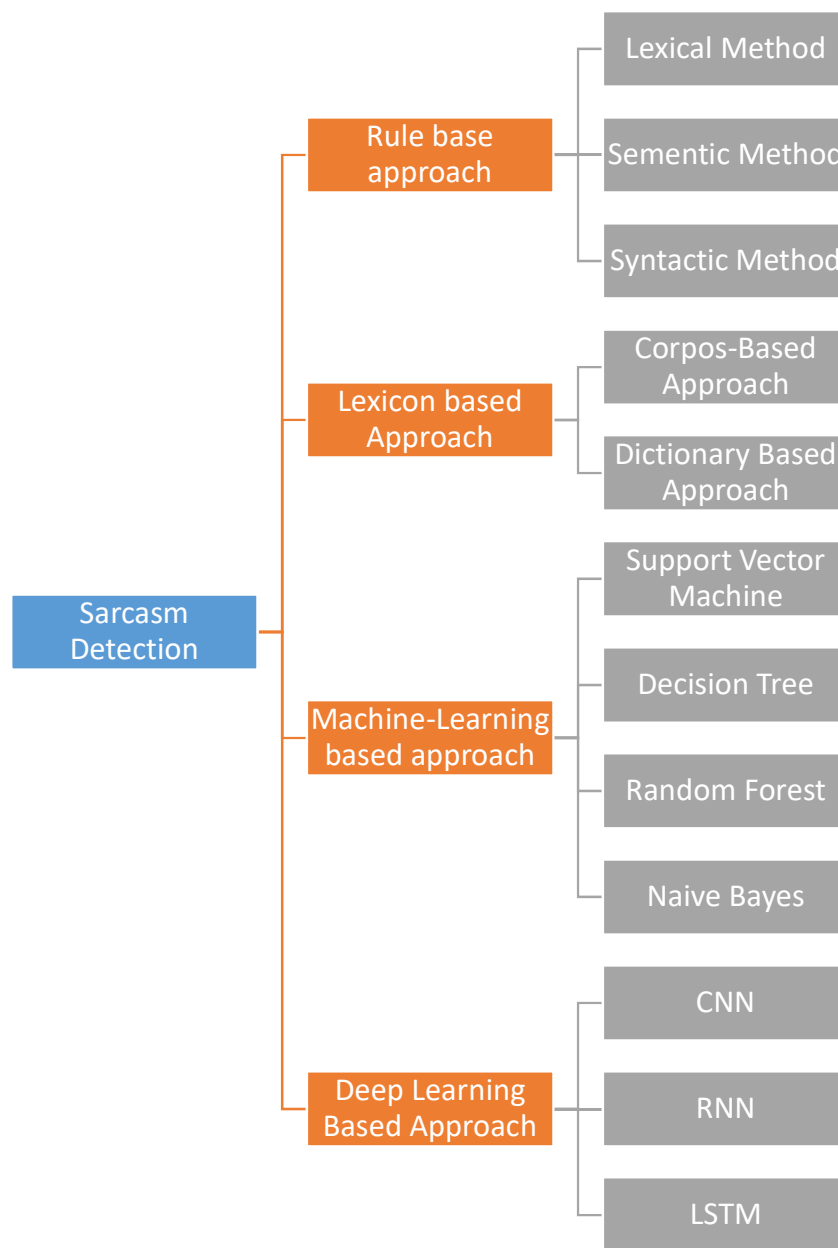


Fig (3) Different methods of Sarcasm Detection

- 1) Rule Based Approach: Using specific information that happens in context and following particular rules, the rule-based approach can identify sarcasm. These rules largely consist of the hashtags #sarcasm, #annoyed and so forth. For instance: "Her performance was incredible!!!" #sarcastic ^[9].

Diana Maynard and Mark Greenwood ^[10] has performed an analysis of the effect of sarcasm scope on the polarity of tweets, and have compiled a number of rules which enable to improve the accuracy of sentiment analysis when sarcasm is known to be present. They considered in particular the effect of sentiment and sarcasm contained in hashtags, and have developed a hashtag tokeniser for GATE, so that sentiment and sarcasm found within hashtags can be detected more easily.

Kreuz and Caucci ^[11] studied the role that different lexical factors play, such as interjections (e.g., “gee” or “gosh”) and punctuation symbols (e.g., “?”) in recognizing sarcasm in narratives.

1) Lexicon-based approach :

The main concept of the lexicon-based approach is to use different opinion words to represent different sentiments. A lexicon is a collection of words or a dictionary of a language. Peter clews and Dr,Joanne Kuzma ^[12] has used lexicon based method to detect sarcasm. In their study, they applied string matching against positive sentiment and interjection lexicons to test if the presence of both can be used to classify content as being sarcastic. The result shows that the most frequently used terms are too generic to be suitable for a sarcasm specific lexicon. It further shows that Boolean matches to two lexicons are suitable for classification of text as being sarcastic

2) Machine Learning based approach :

Machine learning is branch of AI. In machine learning, machine is trained, learned and adapted to work without human assistance. Support Vector Machine, Naïve Bayes, Random Forest, Decision Tree are some of the example of machine learning algorithm which can be used for detecting sarcasm.

Vaishvi prayag jariwala ^[13] has used Support Vector Machine algorithm for News sarcasm detection. Result shows that SVM perform better when the optimal feature set is given as an input. The provided classification system can be improved further by using various other features such as word embeddings, POS tags and other language based features.

3) Deep Learning Based approach :

Deep learning is a type of machine learning and artificial intelligence (AI) that imitates the way humans gain certain types of knowledge. The building block of deep learning is Artificial Neural Network (ANN).Deep learning method has many neurons and complex methods of connecting features. The main type of deep learning methods are CNN (Convolutional neural network), RNN (Recurrent neural network) and Long short term Memory (LSTM).

Sayad Salim et al. ^[14] has used Deep LSTM-RNN with word embedding for detecting sarcasm in Twitter. They achieved 88% of accuracy.

Mayur et al. ^[15] proposed a unique deep neural network model whose Bidirectional LSTM undergo Hyper parameters optimization using genetic algorithm followed by a Convolution Neural Network for sarcasm detection. The proposed approach would perform better than the state of the art models, as they rely mostly on the lexical features. BERT provides a way to interpret sarcasm in a contextual way. Contextualized information provides a better hindsight to get in-depth understanding of text. The proposed model would give accuracy around 93- 95% due to the parameter optimization.

All the four techniques of Sarcasm detection are compared below in the table no.2

Author	Proposed Approach	Datasets	Results
Diana Maynard, Mark A. Greenwood	Rule-Based approach using Hash-tag method	Twitter dataset	Precision: 91%

Vaishvi prayag jariwala[10]	Support Vector Machine	News headline	Precision:74.52 % Accuracy :78.82%
Salim and Agrawal[11]	Deep LSTM-RNN	Twitter	Accuracy : 85%

Table 2. Comparison of sarcasm detection method

VI. Major Issue in Sarcasm Detection

The performance of the sentiment analysis can be improved if the sarcastic words are identified correctly. Some of the main challenge in sarcasm detection are:

- Understanding the ambiguity of the words and correctly interpret the inner meaning of words and classify them as sarcastic or non-sarcastic.
- Many features are available to consider during sarcasm detection, selecting appropriate features and feature extraction technique is also a challenging task.
- Another Challenge is the quality of the data, as the data is in unstructured format, it contain noisy data. How to prepare and pre-process the data before detecting a sarcasm is also a challenge.
- Choosing the proper classification model for sarcasm detection is the major challenge. Some models may work more accurately for one data set may not work for another dataset.

VII. Conclusion and Future Work:

Sarcasm Detection is emerging field in computer science and has grown significantly in the past few years. In this paper, we have given details on types of sarcasm, steps in sarcasm detection and various approaches for sarcasm detection along with its major issue and challenges. Our future work will try to implement sarcasm detection in some dataset and compare various techniques of sarcasm detection on this dataset and compare which one is better to detect sarcasm.

VIII. References:

- <https://www.oxfordlearnersdictionaries.com/definition/english/sarcasm.>
- <https://dictionary.cambridge.org/dictionary/english/sarcasm.>
- <https://www.collinsdictionary.com/dictionary/english/sarcasm.>
- Rajadesingan, A., Zafarani, R. and Liu, H., 2015, February. Sarcasm detection on twitter: A behavioral modeling approach. In *Proceedings of the eighth ACM international conference on web search and data mining* (pp. 97-106).
- Bharti, S.K., Babu, K.S. and Jena, S.K., 2015, August. Parsing-based sarcasm sentiment recognition in twitter data. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (pp. 1373-1380).
- Camp, E., 2012. Sarcasm, pretense, and the semantics/pragmatics distinction. *Noûs*, 46(4), pp.587-634.
- Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N. and Huang, R., 2013, October. Sarcasm as contrast between a positive sentiment and negative situation. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 704-714).

- [8] ABOOBAKER, J. and Ilavarasan, E., 2020. "A Survey on Sarcasm Detection Approaches". *Indian Journal of Computer Science and Engineering*, 11, pp.751-771.
- [9] Bagate, R.A. and Suguna, R., 2019, September. "Different Approaches in Sarcasm Detection: A Survey". In *International Conference on Intelligent Data Communication Technologies and Internet of Things* (pp. 425-433). Springer, Cham.
- [10] Maynard, D.G. and Greenwood, M.A., 2014, March. "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis". In *Lrec 2014 proceedings*. ELRA.
- [11] Kreuz, R. and Caucci, G., 2007, April. "Lexical influences on the perception of sarcasm". In *Proceedings of the Workshop on computational approaches to Figurative Language* (pp. 1-4).
- [12] Clews, P. and Kuzma, J., 2017. "Rudimentary lexicon based method for sarcasm detection". *International Journal of Academic Research and Reflection*, 5(4), pp.24-33.
- [13] Jariwala, V.P., 2020. "Optimal Feature Extraction based Machine Learning Approach for Sarcasm Type Detection in News Headlines". *International Journal of Computer Applications*, 975, p.8887.
- [14] Salim, S.S., Ghanshyam, A.N., Ashok, D.M., Mazahir, D.B. and Thakare, B.S., 2020, June." Deep LSTM-RNN with word embedding for sarcasm detection on Twitter". In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-4). IEEE.
- [15] Ashok, D.M., Ghanshyam, A.N., Salim, S.S., Mazahir, D.B. and Thakare, B.S., 2020, June. "Sarcasm Detection using Genetic Optimization on LSTM with CNN". In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-4). IEEE.