International Information and Engineering Technology Association
Advancing the World of Information and Engineering

International Journal of Computational Methods and Experimental Measurements
Vol. 12, No. 4, December, 2024, pp. 411-420

Journal homepage: http://iieta.org/journals/ijcmem

Enhanced Sarcasm Detection in Telugu Dialogue Systems Using Self Attention-Based RNN and Gated Recurrent Unit Models

Bhargavi Vemala, M. Humera Khanam*

Department of Computer Science and Engineering, Sri Venkateswara University, Tirupati 517502, India

Corresponding Author Email: humera.svec@gmail.com

https://doi.org/10.18280/ijcmem.120410

Received: 5 October 2024
Revised: 29 November 2024
Accepted: 7 December 2024
Available online: 27 December 2024

Keywords:
deep learning, GloVe, GRU, low-resource language, natural language processing, sarcasm detection, self-attention based RNN, Telugu

ABSTRACT

Sarcasm detection is challenging in sentiment analysis, especially for morphologically complex languages like Telugu. Sarcastic statements often use positive words to convey negative sentiments, complicating automated interpretation. Existing sarcasm detection systems predominantly cater to English, leaving a gap for low-resource languages such as Hindi, Telugu, Tamil, Arabic, and others. This study fills this gap by creating and annotating a Telugu conversational dataset, which includes both standard and sarcastic responses. We employed two deep learning models—Self Attention-based Recurrent Neural Network (SA-RNN) and Gated Recurrent Unit (GRU)—to analyze this dataset. Results showed that the SA-RNN model outperformed the GRU, achieving 96% accuracy compared to 94%. The models utilized GloVe word embeddings and specific linguistic features, such as interjections and punctuation marks, to improve sarcasm detection. This research advances the field of sarcasm detection for low-resource languages, particularly Telugu.

1. INTRODUCTION

Sarcasm detection is a complex yet essential aspect of sentiment analysis, particularly in the context of morphologically rich languages such as Telugu. Sentiment analysis, which involves the evaluation of opinions, emotions, and sentiments towards various targets, is significantly challenged by the presence of sarcasm. Sarcasm often employs positive language to convey negative sentiments, thereby complicating the task of automated sentiment detection systems. This linguistic phenomenon is not merely a stylistic choice; it is a deliberate form of communication that can obscure the true intent behind a statement, making it critical for sentiment analysis systems to accurately identify sarcasm to enhance their effectiveness [1]. Sarcasm detection is a critical aspect of sentiment analysis, which seeks to interpret the underlying emotions and opinions expressed in textual data. This nuanced form of communication often involves stating the opposite of what is meant, thereby complicating the sentiment analysis process. Sarcasm can flip the sentiment value of a statement, making it challenging for automated systems to accurately gauge the intended sentiment. For instance, a seemingly positive statement such as "I love being ignored" conveys a negative sentiment due to the context of being ignored, illustrating the complexity inherent in sarcasm detection [2].

2. RELATED WORK

The intricacies of sarcasm are further compounded in low-resource languages, where existing models and datasets predominantly cater to English and a few other widely spoken languages. Examine that while there has been substantial progress in sarcasm detection in English, the same cannot be said for languages like Hindi, Telugu, and Tamil, which are rich in morphology and often employ unique linguistic features [3, 4]. The scarcity of annotated datasets in these languages presents a significant barrier to developing effective sarcasm detection algorithms. For instance, Telugu, the second most spoken language in India, has a rich literary heritage and a growing presence on social media, yet it remains underrepresented in the field of natural language processing (NLP) [4]. The challenge is exacerbated in languages that are morphologically rich, such as Telugu, where the structure and form of words can significantly alter meaning and sentiment interpretation. To address this gap, recent studies have focused on creating annotated corpora specifically for Telugu, utilizing conversational data from popular comedy shows. This approach not only provides a rich source of sarcastic dialogue but also aligns with the cultural context in which sarcasm is frequently employed. By analyzing these conversational patterns, researchers have developed algorithms that leverage hyperbolic features—such as interjections and intensifiers—to improve sarcasm detection accuracy [1]. The application of deep learning models, particularly attention-based Recurrent Neural Networks and Graph RNN, has shown promising results, with the latter achieving higher accuracy in sarcasm detection tasks [5]. The detection of sarcasm in Telugu presents a unique set of challenges and opportunities. By focusing on the development of tailored datasets and advanced deep learning models, investigators aim to enhance the accuracy of sentiment analysis systems in low-resource languages. This work not only contributes to the field of NLP but also fosters a deeper understanding of the linguistic nuances that characterize human communication in diverse cultural contexts. various methodologies and previous research outlined in study [2].

The prevalence of sarcasm in everyday communication highlights its importance in sentiment analysis, particularly in social media contexts where informal language and cultural nuances play a significant role. Sarcasm is often intentional and can serve various communicative functions, including humor, criticism, or social bonding. Absence of vocal variation and facial cues in written text makes it difficult for algorithms to detect sarcasm reliably. Existing sarcasm detection models have predominantly focused on English, leaving a significant gap in research for low-resource languages, including Hindi, Telugu, Tamil, and others [6]. This gap is particularly concerning given the increasing use of regional languages in digital communication, where sarcasm is frequently employed. In the context of Indian languages, Telugu stands out due to its rich morphological structure and cultural significance. It ranks as the second most widely spoken language in India, and its speakers increasingly engage in social media communication in their native tongue. This shift necessitates the development of automated sentiment analysis tools that can effectively handle sarcasm in Telugu. The morphological richness of Telugu presents unique challenges for researchers, as traditional sentiment analysis techniques may not be directly applicable [7, 8]. The existing literature indicates that while there have been efforts to address sarcasm detection in other languages, the specific challenges posed by Telugu remain largely unexamined.

To address this gap, recent studies have focused on collecting and annotating corpora of Telugu conversational sentences, particularly those derived from comedy shows where sarcasm is prevalent. These datasets typically consist of exchanges that feature a question followed by a sarcastic response, providing a rich source of data for training sarcasm detection models. The development of algorithms that leverage hyperbolic features—such as interjections, intensifiers, and punctuation—has shown promise in enhancing the accuracy of sarcasm detection in Telugu [9, 10]. The application of deep learning models, including (ABRNNs) and RNNs, has further advanced the field, demonstrating improved performance in recognizing sarcastic utterances compared to traditional methods [11].

The significance of sarcasm detection extends beyond mere sentiment analysis; it has implications for various applications, including brand management, customer feedback analysis, and social media monitoring. As sarcasm can distort the perceived sentiment of a message, accurately identifying it is crucial for businesses and organizations seeking to understand public opinion and sentiment accurately. Moreover, the ability to detect sarcasm can enhance the performance of dialogue systems and conversational agents, allowing them to respond more appropriately to user inputs [12, 13]. In the detection of sarcasm in Telugu and other low-resource languages presents both challenges and opportunities for researchers in the field of natural language processing. The development of specialized algorithms and the collection of annotated datasets are essential steps toward improving sarcasm detection capabilities. As the digital landscape continues to evolve, the need for effective sentiment analysis tools that can navigate the complexities of sarcasm will only grow, underscoring the importance of this research area [14, 15].

2.1 Sarcasm detection

Linguistic characteristics are essential for identifying irony and sarcasm in written language, particularly in the context of sentiment analysis. In the realm of sarcasm detection, various lexical and syntactic features have been employed to enhance the accuracy of identifying sarcastic expressions. For instance, a semi-supervised approach has been employed to identify sarcasm in tweets and Amazon product reviews, utilizing two key lexical features: pattern and punctuation-based features. These features contribute to the development of a weighted (KNN) categorization representation for effective sarcasm detection [16, 17]. In addition to lexical features, numerous other linguistic indicators have been explored to identify sarcasm in tweets. These include attributes obtained from linguistic analysis and word frequency, as well as WordNet affect, which provides a semantic framework for understanding emotional connotations. Pragmatic features, and replies, have also been integrated into sarcasm detection frameworks, emphasizing the multifaceted nature of sarcasm as a communicative phenomenon [18]. A well-constructed lexicon-based approach has been proposed, based on the premise that sarcastic tweets often present a contrast between a positive sentiment and a negative situation. This approach utilizes unigram, bigram, and trigram features for lexicon generation, enhancing the model's ability to detect sarcasm effectively.

Hyperbolic features, such as intensifiers and interjections, have been identified as particularly effective in sarcasm detection. For example, the utterance "fantastic weather when it rains" is more readily recognized as sarcastic compared to a non-hyperbolic statement. The presence of hyperbolic language serves as a strong indicator of sarcasm, allowing for more accurate interpretation of the speaker's intent. Furthermore, parsing techniques have been employed to segment tweets into phrases, facilitating the generation of lexicon files that support sarcasm identification in Twitter data. Beyond lexical and syntactic features, the behavioral traits of Twitter users have emerged as valuable indicators for sarcasm detection. The contextual nature of sarcasm necessitates a shared understanding between the speaker and the audience, which can be informed by the user's past behavior and sentiment expressed in previous tweets. This behavioral context has been integrated into sarcasm detection frameworks, drawing on theories drawn from behavioural and psychological research to develop a comprehensive modeling strategy [19]. Moreover, a system has been developed to identify sarcastic tweets in the context of predicting electoral outcomes by analyzing public sentiment on Twitter. This system utilizes various features, including exclamation marks, question marks, hashtags, emoticons, adjectives, and verbs, to ascertain the sarcastic polarity of tweets through a supervised deep learning approach [2]. The incorporation of researchers early tweets provides additional context for sarcasm detection, enhancing the model's predictive capabilities [1]. In the detection of sarcasm in text, particularly in languages like Telugu, requires a comprehensive approach that integrates lexical, syntactic, and behavioral features. The development of advanced frameworks that consider these diverse features is essential for improving the accuracy of sarcasm detection in dialogue systems. This research contributes to the ongoing efforts to enhance sentiment analysis capabilities, particularly in low-resource languages, by leveraging deep learning models such as attention-based RNN and graph RNN to analyze the complexities of sarcastic expressions in Telugu conversational data [20, 21].

2.2 Irony exposure in LRL

The initial work on irony exposure in LRL was conducted on Indonesian social media data [22]. This dataset is physically collected from Twitter and introduced two additional features for irony detection, beyond standard emotion analysis: negativity information and count of interjection words. The study also utilized a transformed version of SentiWordNet for emotion categorization. Thelwall et al. [23] highlighted the vast no. of casual messages posted daily on social networks, blogs, and discussion forums. Existing algorithms have been designed to identify both sentiment and sentiment strength, aiding in understanding emotional expression in informal communication and detecting inappropriate or anomalous affective utterances, which could indicate potentially harmful behavior. A set of characters precisely aimed at detecting irony in social media was developed, along with a novel (MSELA) to address the class imbalance problem in English and Chinese judgments [24]. Additionally, a system is proposed for detecting sarcastic sentences in Hindi using support [25], focusing on features such as emoticons and punctuation marks. To the best of our knowledge, no prior research has been conducted on sarcasm detection in the Telugu language.

3. PROPOSED METHODOLOGY

The proposed scheme for irony recognition in Telugu conversational sentences involves a multi-faceted approach utilizing deep learning techniques. This scheme integrates the following components:

Data Collection and Annotation: We have gathered and annotated an extensive corpus of Telugu conversational sentences, structured with two distinct conditions: (1) a usual question normal reply, and (2) a usual question followed by ironic reply.

Feature Extraction: The GloVe methodology builds word embeddings by capturing word co-occurrence information from a large corpus. It creates a co-occurrence matrix, minimizes a log-probability-based objective function, and uses gradient descent to learn word vectors. These vectors are then used to capture semantic similarities, making GloVe a powerful tool for representing words in a meaningful way for a wide range of NLP tasks.

Deep Learning Models: Two advanced deep learning models are employed:

SA-RNN: This model focuses on identifying contextual relationships within the conversation to detect sarcasm.

GRU: The GRU is a type of RNN architecture designed to solve some of the limitations found in traditional RNNs, particularly the issues of vanishing and exploding gradients, which hinder the ability to model long-term dependencies in sequential data. The GRU is an improvement over the standard RNN and is a simplified version of the Long Short-Term Memory (LSTM) network, making it computationally efficient while retaining effectiveness.

Evaluation and Comparison: The performance of both models is evaluated using accuracy metrics. Preliminary results indicate that the GRU model outperforms the SA-RNN, achieving an accuracy of 96% compared to 94%.

This proposed scheme aims to advance sarcasm detection capabilities in Telugu by leveraging state-of-the-art deep learning techniques and feature extraction methods tailored to the unique characteristics of the language.

3.1 Sarcasm detection model

The demonstrate for irony exposure, detailing steps of Telugu data gathering and explanation, followed by (POS) labeling and investigation to develop rules for identifying irony in Telugu conversational sentences.

3.2 Comprehensive data collection process

The task of irony detection in Telugu dialogue systems is particularly challenging due to the language's morphological richness and the scarcity of annotated datasets. To address this issue, we undertook a comprehensive data collection effort, manually gathering approximately 6900 conversational sentences from various sources, including popular Telugu TV comedy shows. This involved watching nearly 480 archived episodes to extract relevant dialogues, which were structured primarily as questions followed by replies. The gathered dataset has been made openly accessible on GitHub, facilitating further research in this area. The annotated corpus is designed to support the development of sarcasm detection models, particularly focusing on two conditions: normal questions with normal replies and normal questions with sarcastic replies. To analyze this dataset, we employed two advanced deep learning models: SA-RNN and GRU. Our findings indicate that the Graph RNN model significantly outperformed the SA-RNN, achieving an accuracy. By leveraging hyperbolic features such as interjections, intensifiers, question marks, and exclamation symbols, our models enhance the detection of sarcasm in Telugu conversational sentences. This research not only contributes to the field of sarcasm detection for low-resource languages but also underscores the importance of developing tailored approaches for languages like Telugu, where existing resources are limited.

3.3 Data classification

The dataset for this study is distributed among Telugu language experts, including teachers and practitioners, who provided valuable feedback by manually annotating 6900 sentences to determine whether each sentence is satirical or not. Subsequently gathering annotations from three participants, we observed these conversational sentences followed one of two patterns (1) a usual question normal reply, or (2) Usual question followed by ironic reply. Tables 1 and 2 show a sample of POS and annotated conversation sentences, one for each of these patterns. The annotation process revealed that most sentences fell into the second pattern, where a usual question followed by ironic reply.

To assess (IAA), we used two coefficients: Cohen's Kappa [26] and Fleiss Kappa [27]. When there are two annotators, Cohen's Kappa is used, but for more than two annotators, Fleiss Kappa is appropriate. Since this study involved three annotators, we used the Fleiss Kappa coefficient, calculated using the formula shown in Eq. (1). The IAA for this study is

0.85, indicating a high level of agreement among the annotators.

Table 1. List of POS tag set used in this work

No.	Category	POS Tag	Example
1	Noun	NN	పుస్తకం (pusthakam), కార్తీక్ (Karthik)
2	Pronoun	PRP	అతడు (athadu), ఇది (idi)
3	Verb	VM	చదవు (chaduvu), వ్రాయు (vrayu)
4	Adjective	JJ	తెల్ల (Tella), చక్కని (chakkani)
5	Adverb	RB	త్వరగా (Tvaraga), అక్కడ (Akkada)
6	Conjunction	CC	మరియు (Mariyu), అయితే (Ayithe)
7	Interjection	INJ	ఛీ (Chi), అహో (Aha), అలానా (Alana)
8	Numerical	NUM	ఒకటి (Okaṭi), రెండు (Reṇḍu)
9	Determiner	DEM	ఆ (A), ఏ (Ē)
10	Question Words	WQ	ఎక్కడ (Ekkada), ఏం (Ēṃ), ఎలా (ela)
11	Symbol	SYM	. , ? !

Table 2. Explained (TSC)

Telugu Conversation	English Meaning	Annotation
నెల క్రితమేగా పెళ్లైంది, అంతలోనే విడాకులు కావాలంటున్నావు దేనికే, ఏమైనా కొడుతున్నాడా? లేదు నాకు వంట చేయడం రాదని నిన్ననే చెప్పాడు అందుకే విడాకులు అడుగుతున్నాను!	"You got married just a month ago, and now you're asking for a divorce? Why? Is he hitting you ?" "No, he just told me yesterday that he doesn't know how to cook. That's why I'm asking for a divorce!"	Sarcastic
ఏమే సుమ నాకు 1000 రూపాయలు దొరికాయే. అవునా అయితే 50 50 షేర్ చేసుకుందాం. హా మరి మిగిలిన 900 రూపాయలు ఎవరికీ ఇద్దామే?	"Hey Suma, I found 1000 rupees!" Then let's share it 50-50." "Okay, but to whom should we give the remaining 900 rupees?"	Sarcastic
ఇదిగో ఆఖరి సారిగా చెబుతున్నా. మీ తలమీద వెంట్రుకలు ఇప్పటికే చాలా రాలి పోయాయి. ఇదే ఇంకా కొనసాగితే మిమల్ని వదిలేసి పుట్టింటికి వెళ్ళిపోతా. ఛ, ఛ. జుట్టు రాలి పోతుంది అని అనవసరంగా ఇన్ని రోజులు తెగ బాధ పడిపోయాను. ఈ విష్యం ముందుగా తెలిసుంటే జుట్టు రాలిపోవడం కోసం అసలు పట్టించుకునే వాన్నే కాదు.	"Look, I'm telling you for the last time. You've already lost a lot of hair. If this continues, I'll leave you and go back to my parents' home." "chi, Chi. I was unnecessarily worried all these days thinking about my hair falling out. If I had known this earlier, I wouldn't have even cared about losing hair."	Sarcastic

$$k=P-P_e/1-P_e \quad (1)$$

Our analysis is based on the following assumptions:

- Sentences that follow the pattern "Usual question normal reply" are classified as non-ironic.
- Sentences that follow the pattern "Usual question followed by ironic reply" are considered ironic.

From these statements, we discovered that out of 6900 sentences, 6400 were interpretedas sarcastic, and the remaining 500 were non-sarcastic.

In this study, we focused on sentences that adhere to "Usual question followed by ironic reply" pattern, as this pattern is prevalent, comprising approximately 97% of the sarcastic sentences in the dataset.

Following annotation, 6400 out of 6900 sentences were identified as sarcastic. Of the 6400 sarcastic sentences, 6244 (97%) tracked sample of a Usual question followed by ironic reply. All 6244 ironic sentences were used to develop the rules for sarcasm detection.

A separate testing set comprising 1380 sentences is utilized, which is not part of the original dataset. Figure 1 shows the recognition model of conversational sentences.

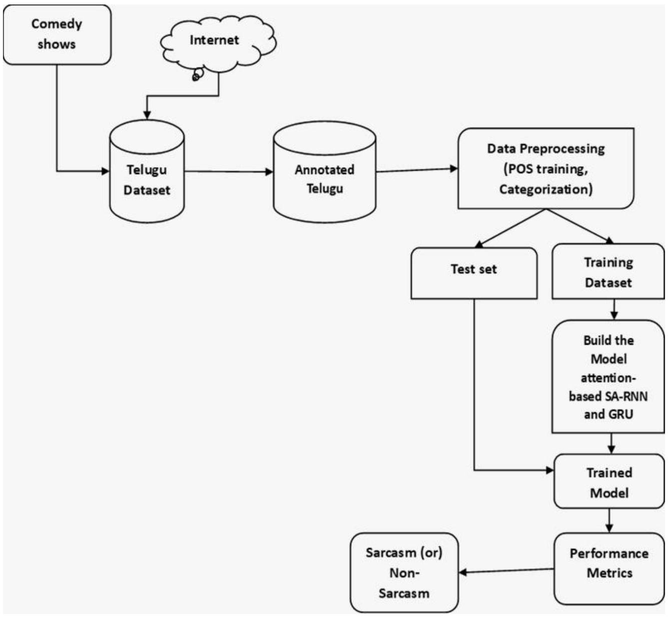

Figure 1. Irony recognition model

3.4 Morphosyntactic labeling (POS)

Part-of-Speech (POS) tagging is process of assigning the appropriate grammatical category to every word in a performed sentence. POS taggers are obtained by modeling morphosyntactic structure of natural language processing. Telugu POS tagging program used in this study is related to Model 5 labelled in Table 1 of study [28], but specifically tailored for Telugu. The corpora for this tagger were downloaded, cleaned, and tagged using a high-precision, low-recall approach. Due to extensive training on substantial datasets, tagger is capable of handling a broad vocabulary and predicting tags of unidentified words based on identified words. The tagging process employs a Hidden Markov Model (HMM) approach and utilizes the standard Indian language tag set [29], which consists of 21 tags. The Telugu tagger utilized in this study is based on the TnT tagger, renowned for its robustness and speed.

Table 2 provides examples of some of the Telugu tags used in this study, and Table 3 shows examples of Telugu sentences along with their associated Part-of-Speech (POS) tag information.

Table 3. Example of sarcasm data in Telugu sentences

Telugu Conversation	English Meaning	Annotation
ఏవండి మీకోసం ఇష్టం అని బిర్యానీ చేశాను ఎలా ఉంది నా వంట? చాల బాగుంది బిర్యానీ రాజ్యం. ఎక్కడ నేర్చుకున్నావు? యూట్యూబ్లో చూసి నేర్చుకున్నండి.	"Hey, I made biryani because I know you like it. How is my cooking?" "It's really good, biryani queen! Where did you learn to cook like this?" "I learned it by watching YouTube."	Non-sarcastic
ఏవండి ఇవాళ కొత్త సినిమా ఏదయినా రిలీజ్ అయిందా, సినిమాకి వెళ్దాం. ఏం అయినట్టు లేదు, వచ్చే వారం రిలీజ్ అవుతుంది, వెళ్దాం.	"Hey, is there any new movie releasing today? Let's go to the movies." "Nothing like that. It will release next week. Let's go."	Non-sarcastic

3.5 Recommended set of rules

In this study, we analyzed 6400 sarcastic sentences from the training set and categorized them into two distinct types:

Sentences that begin with a TNW.

Sentences that begin with a TIW.

Tables 4 and 5 provide lists of TNW and TIW, respectively. Based on our observations of sarcastic sentences in Telugu conversations, we propose three algorithms to detect sarcasm in these categories. TNWS, TIWS and Responding to a statement by rephrasing it as a question.

Table 4. POS labelled

Telugu Sentence	Parts-of-Speech Tagging
నెల క్రితమేగా పెళ్లైంది, అంతలోనే విడాకులు కావాలంటున్నావు దేనికే, ఏమైనా కొడుతున్నాడా? లేదు నాకు వంట చేయడం రాదని నిన్ననే చెప్పాడు అందుకే విడాకులు అడుగుతున్నాను!	NN RB PP SYM VM RB NN MV PRP SYM PRP VM SYM AUX PRP NN VM VM RB VM SYM
ఏవండి ఇవాళ కొత్త సినిమా ఏదయినా రిలీజ్ అయిందా, సినిమాకి వెళ్దాం. ఏం అయినట్టు లేదు, వచ్చే వారం రిలీజ్ అవుతుంది, వెళ్దాం.	INJ NN JJ NN PRP VM VM SYM NN PRP VM SYM PRP CC PRP SYM JJ NN VM VM SYM VM SYM

Table 5. List of TNW

Negation Word	English Meaning
లేదు (Ledu)	No, Not, does not exist
అలా కాదు (Ala Kadu)	Not like that
కాదు (Kadu)	Not
కాకుండా (kakunda)	Without
వద్దు (Vaddu)	Don't want

3.6 Telugu_Negation_Prefix_Algorithm

The TNWS algorithm is designed to discover irony in Telugu conversational sentences by leveraging specific linguistic patterns. The algorithm identifies sarcasm based on the occurrence of TNW such as "ledu," "kaadu," and "vaddu" at beginning of a sentence, combined with the presence of punctuation marks like question marks or exclamation marks at the end. The algorithm first extracts the first and last words of each sentence, then checks if the first word is one of the predefined negation words and if the last word matches any of the specified punctuation marks. If both conditions are satisfied, sentence is categorized as ironic; otherwise, it is not. This method is implemented in Python, where a function processes a list of sentences, classifying each based on these criteria and providing a straightforward classification result.

3.7 TIWS word

This algorithm focuses on detecting sarcasm in Telugu sentences that begin with interjection words like "ayyo," "haa," or "alaana." During analysis of Telugu ironic replies, it is examined that many sarcastic responses started with these interjection words. Based on this observation, Table 6 interjection words and its examples shown in Table 7.

Table 6. List of Telugu interjection words

Interjection Word	English Meaning
ఓహో (Oho)	Oh! , I See!
ఛీ (Chi)	Yuck!
అయ్యో (Ayyo)	Oh No!
అయ్య బాబోయ్ (Ayya Baboy)	Oh my God!
అమ్మో (Ammo)	Oh no! / Oh my!
వహ్ (Vah)	Great! / Fantastic!
హా (Haa)	Ha!
అలానా (Alana)	"Is it so?", Really?
ఆహా (Aha)	Aha! Wow!

GloVe embeddings are utilized for the linguistic elements pertinent to sarcasm, including interjections, punctuation, and ellipses. Although GloVe embeddings are robust, they are static and may not entirely encapsulate the contextual significance of words. pre-trained model IndicBERT is a transformer-based language model accessible on GitHub that captures subtle linguistic patterns and context-dependent sarcasm cues.

Fuse GloVe embeddings plus IndicBERT embeddings for each word as features to be fed to the model. Utilize GloVe features with GRUs to capture sequential patterns, and employ IndicBERT for transformer-based contextual comprehension. Integrate characteristics from both embeddings through concatenation methods. Input the integrated characteristics into fully connected layers for sarcastic classification.

In comparison to attention-based RNN and GRU models, the rule-based approach leverages explicit linguistic patterns like interjections and POS tags to detect sarcasm, whereas the deep learning models utilize context-rich word embeddings like GloVe to capture sarcasm. The SA-RNN performed better, achieving higher accuracy (96%) due to its capability to weigh important parts of the sentence more effectively compared to the GRU model, which had a slightly lower accuracy (94%).

3.8 GRU

In the study [30], the GRU uses gating mechanisms to control the flow of information and efficiently learn temporal relationships in time-series data, speech, natural language processing, and other sequential data tasks. Here's an in-depth look at how GRUs work and what sets them apart.

Table 7. Telugu conversation sentences

S. No	Question (Conversation)	Reply	
1	బాబు, ఏం చేస్తున్నావు? వీణ వాయిస్తున్నా ఆంటీ. (Babu, what are you doing? playing the veena Aunty.	**అలానా (Alana)** మరి ఎక్కువ వాయించకు బాబు, తీగలు తెగిపోతాయి! (Alana Don't play too much baby, the strings will break!)	Reply Begin with Interjection Word (**అలానా (Alana)**)
2	ఏంట్రా అగ్గిపుల్ల వెలిగించి చూస్తున్నావు, ఏమైనా పోయిందా? (Why are you lighting matches, is there anything missing?)	**లేదు(Ledu)** కొవ్వొత్తి వెలిగే ఉందో లేదో చూస్తున్నాను అంతే! (No I see if the candle is lightening or not that's it!)	Reply Begin with Negation Word **లేదు(Ledu)**
3	నాకు చీమలను చూస్తే భయమేస్తున్నది డాక్టర్. ఇలా ఎప్పటినుంచి జరుగుతున్నది? (I am Scaring seeing ants Doctor. Since when is this happening?)	**హా (Ha)** నాకు షుగర్ వుందని పరీక్షలో తేలినప్పటి నుండి! (Ha Ever since I tested positive for diabetes!)	Reply Begin with Interjection Word. **హా (Ha)**
4	సార్, మా దొంగల ఫోటోలూ, లిస్టు వుంటే ఇస్తారా? ఎందుకూ? (Sir, can you give us the photos and list of thieves? Why?	**హా(Ha)** మా అమ్మాయికి పెళ్లి చేయాలి, మంచి దొంగోడిని చూద్దామని! (Ha, my girl needs to get married, let's find a good thief!	Reply Begin with Interjection Word. **హా (Ha)**

GRU Structure

A GRU unit consists of two main gates:

Update Gate: Controls how much of the previous information needs to be passed to the next time step.

Reset Gate: Determines how much of the past information to forget.

These gates allow the GRU to adaptively capture dependencies in sequential data without suffering from long-term memory decay, while also avoiding over-complexity by having fewer parameters than LSTM networks.

Update Gate

The update gate controls how much of the current hidden state should be updated with new information and how much should remain from the previous hidden state.

It allows the model to carry forward important past information while integrating new input.

Current Memory Content:

This intermediate hidden state candidate, denoted by h is computed using the reset gate. If the reset gate outputs a value close to 0, it allows the network to ignore the past hidden state completely and focus only on the current input.

Final Hidden State:

The final hidden state for the current time step is computed using a weighted combination of the previous hidden state and the candidate hidden state. The update gate decides how much of the candidate state to combine with the previous hidden state.

4. EXPERIMENTAL RESULTS

4.1 Statistical evaluation metrics for deep learning models

To assess the performance of the proposed deep learning models, SA-RNN and GRU—three statistical evaluation metrics were utilized: Precision, Recall, and F-Score. These metrics provide a comprehensive understanding of how well the models performed in detecting sarcasm.

Precision measures the proportion of correctly identified relevant instances out of all instances predicted as relevant. It indicates how accurate the model is in classifying sarcastic sentences. Recall quantifies the ability of the model to identify all relevant instances, showing how well the model captures sarcastic sentences from the entire dataset. F-Score is the harmonic mean of Precision and Recall, offering a balanced measure that considers both accuracy and completeness in the model's predictions.

The formulas for these metrics are as follows:

Precision=$T_P/T_P + F_P$

Recall = $\frac{T_P}{T_P+F_N}$

F-Score = 2 × Precision × Recall / Precision + Recall

where:

T_P = (True Positive) refers to the number of correctly classified sarcastic sentences.

F_P = (False Positive) refers to the number of non-sarcastic sentences incorrectly classified as sarcastic.

F_N = (False Negative) refers to the sarcastic sentences that were incorrectly classified as non-sarcastic. In this study, the SA-RNN outperformed the GRU, achieving an accuracy of 96% compared to GRU's 94%. Both models utilize hybrid mode word embeddings to capture semantic and syntactic nuances, improving the identification of sarcasm in Telugu dialogues.

4.2 Analysis in the deep learning approaches

In this study, we analyzed the effectiveness of two deep learning models—SA-RNN and GRU—for sarcasm detection in Telugu conversation sentences. The features used for training were extracted from the annotated corpus, leveraging word embeddings such as GloVe specific to the special tokens to capture semantic and syntactic nuances.

Interjections were regarded as distinct tokens during tokenization, ensuring that the model accurately captured the embeddings for these words.

In certain instances, incorporated into the vocabulary, and their embeddings were either pre-trained during the model's training phase. Interjections were addressed independently in certain sections of the model, while punctuation marks were regarded as tokens in both the GloVe and IndicBERT embeddings. Both sets of embeddings were amalgamated using a concatenation approach. The integrated embeddings were subsequently processed by the model to identify sarcasm.

The training and testing splits were varied across multiple trials, starting from a 80-20 split to evaluate the performance of the models.

The performance of the SA-RNN and GRU models was tested individually, as well as with combined features. Hyper parameters of the model shown in Table 8. The results are averaged over 5 trials for each split ratio. The highest accuracy of 96% was achieved by the SA-RNN model with an 80-20 training-to-testing ratio, outperforming the GRU model, which reported an accuracy of 94%. Both models showed consistent performance across the different split ratios, with minimal fluctuations in accuracy.

Table 8. Hyperparameter values

Models	Hyper Parameter Values	
SA-RNN and GRU	Batch Size	256
	Epochs	100
	Hidden Layers	16
	Kernal size	5
	Filters	128
	Dropout, Recurrent dropout	0.2
	Learning Rate	0.001
	Activation Function	Relu,tanh
	Dense Layer (Activation Function)	Softmax
	Optimizer	Adam

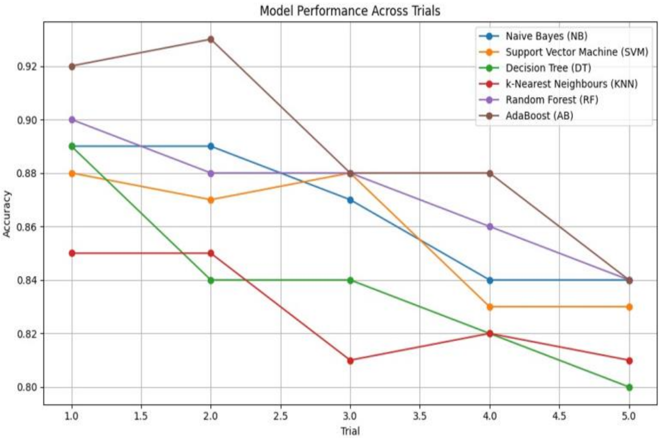

Figure 2. Accuracy of different classifiers of (80-20)

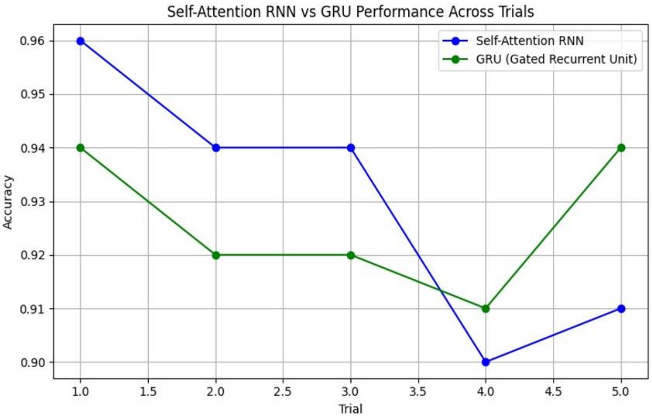

Figure 3. Accuracy of combined algorithm of different classifiers

In these trials, AdaBoost (AB) achieved the highest accuracy of 93% in the second trial, demonstrating its superior performance compared to other classifiers shown in Figure 2.

The experimental results for the SA-RNN model are shown in Figure 3, indicating that it consistently outperformed the GRU model across different trials. In contrast, GRU, although competitive, demonstrated slightly lower accuracy and exhibited minor variations, particularly in certain split ratios. A detailed breakdown of the 5 trials with the 80-20 split ratio is shown in Tables 9 and 10, highlighting the superior performance of the SA-RNN. Figures 4 and 5 illustrate the performance of the GRU model, where the highest accuracy of 94% is observed. However, the GRU model's performance was slightly less stable compared to the SA-RNN.

Table 9. 5 trials of the classifiers with a (80-20) split

Trial	NB	SVM	DT	KNN	RF
1	0.89	0.88	0.89	0.85	0.90
2	0.89	0.87	0.84	0.85	0.88
3	0.87	0.88	0.84	0.81	0.88
4	0.84	0.83	0.82	0.82	0.86
5	0.84	0.83	0.80	0.81	0.84

Notes: NB: Naive Bayes; SVM: Support Vector Machine; DT: Decision Tree; KNN: K-Nearest Neighbors; RF: Random Forest; AB: AdaBoost

Table 10. 5 trails on SA-RNN and GRU (80-20 split-up)

Trail	SA-RNN	GRU
1	0.96	0.94
2	0.94	0.92
3	0.94	0.92
4	0.90	0.91
5	0.91	0.94

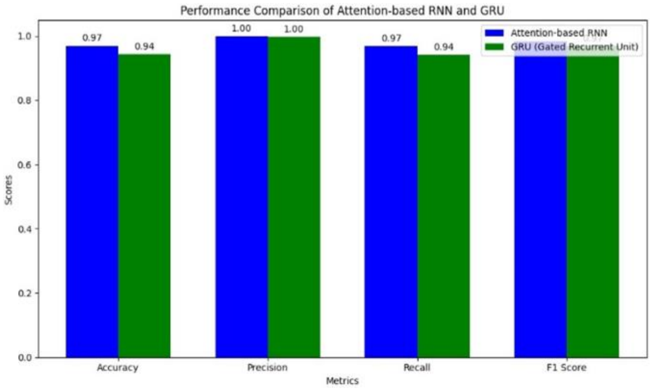

Figure 4. Performance comparision of SA-RNN and GRU

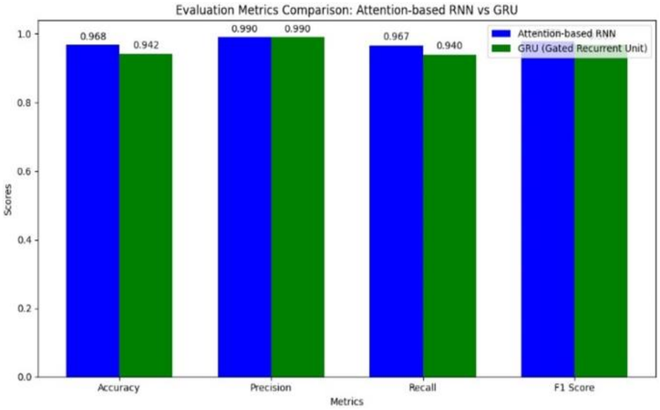

Figure 5. Evalution metric comparision on SA-RNN and GRU

4.3 Experimental evaluation

Experiments is conducted using two deep learning models, SA-RNN and GRU, for sarcasm detection on a test set of 1380 Telugu conversational sentences. The performance of both models was evaluated using a confusion matrix, precision, recall, and F-score. The results of the experimental evaluation for both models are presented in Table 11, with a breakdown of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The combined performance of both models is shown, as well as individual results for each model. Training and validation accuracy shown in Figure 6.

Further evaluation metrics including accuracy, precision, recall, and F-score are provided in Table 12. The SA-RNN model achieved an accuracy of 96%, outperforming the GRU model, which attained an accuracy of 94%. The comprehensive performance of all classifiers is illustrated in Figure 7 and in Table 13. The RNN model showed higher precision, recall, and F-score compared to GRU.

Table 11. Confusion matrix for SA-based RNN and GRU models

Model	TP	FP	TN	FN
SA-RNN	1239	2	98	41
GRU	1204	4	96	76

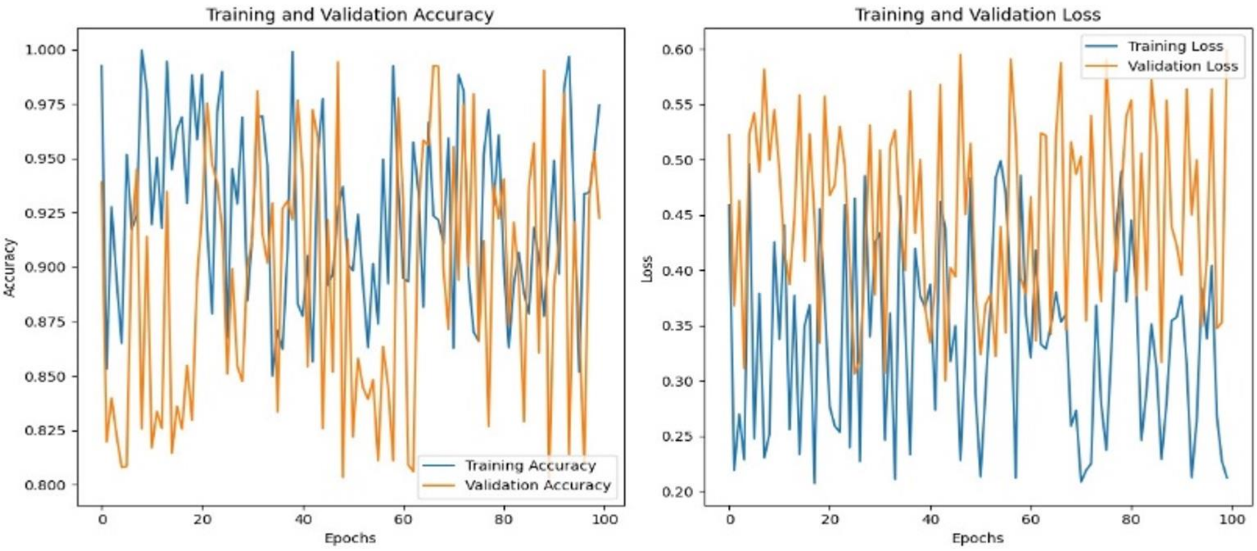

Figure 6. Training performance on models

Table 12. Evaluation metrics for SA-based RNN and GRU models

Model	Accuracy	Precision	Recall	F1-Score
SA-RNN	0.968	0.990	0.967	0.979
GRU	0.942	0.990	0.940	0.969

The results demonstrate that the SA-RNN model, enhanced by GloVe word embeddings, performed better across all metrics. While the GRU model was competitive, it fell slightly behind in both precision and recall. The use of GloVe embeddings played a crucial role in enhancing the models' ability to detect sarcasm in Telugu conversational sentences.

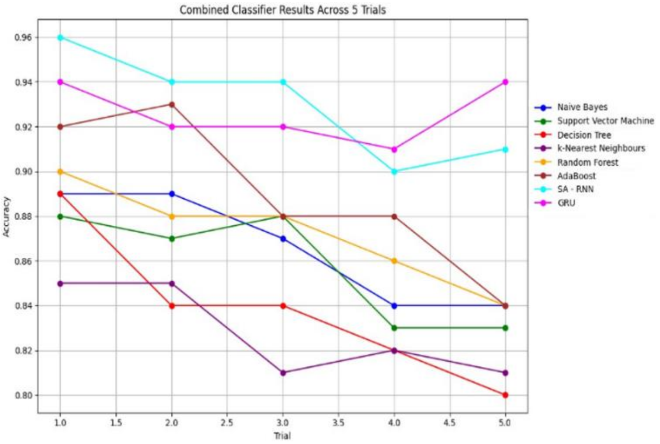

Figure 7. Combined classifier results across 5 trails

Table 13. 5 trials of the classifiers with combined results

Trial	NB	SVM	DT	KNN	RF	AB	SA-RNN	GRU
1	0.89	0.88	0.89	0.85	0.90	0.92	0.96	0.94
2	0.89	0.87	0.84	0.85	0.88	0.93	0.94	0.92
3	0.87	0.88	0.84	0.81	0.88	0.88	0.94	0.92
4	0.84	0.83	0.82	0.82	0.86	0.88	0.90	0.91
5	0.84	0.83	0.80	0.81	0.84	0.84	0.91	0.94

5.CONCLUSION

In the domain of sarcasm detection for low-resource languages such as Hindi, Telugu, Tamil, Arabic, and others, there has been limited work due to the scarcity of datasets for analysis and experimentation. The collection of such datasets is one of the most challenging tasks in this field. In this study, we manually built and annotated a dataset of Telugu conversational sentences from various sources, labeling them as sarcastic or non-sarcastic.

To identify sarcasm in the collected dataset, we applied two deep learning models: SA-RNN and GRU. Both models were trained on the annotated dataset, and we found that the SA-RNN model outperformed the GRU model, achieving an accuracy of 96% compared to 94%. This performance is enhanced by leveraging GloVe word embeddings, which helped capture the nuanced linguistic features essential for sarcasm detection in Telugu conversations.

This work marks a significant step towards improving sarcasm detection in low-resource languages, as there has been no reported work on sarcasm detection specifically for Telugu.

The results demonstrate the potential of deep learning approaches in this area, achieving high accuracy with the limited dataset available. The proposed models can serve as a foundation for future research and development in sarcasm detection for low-resource languages.

Future Scope

There are several promising avenues for further research and improvement in sarcasm detection for low-resource languages like Telugu:

1. **Expansion of Datasets:** The accuracy and robustness of sarcasm detection models can be significantly enhanced by expanding the dataset. Collecting and annotating larger datasets from diverse sources, such as social media, news platforms, and conversation transcripts, would provide a more comprehensive training set, allowing models to generalize better across various contexts and dialects.
2. **Multilingual Sarcasm Detection:** Extending this work to include multiple low-resource languages such as Hindi, Tamil, and Arabic can help develop a more universal model for sarcasm detection. Leveraging transfer learning techniques could allow knowledge from one language to improve performance in another, thereby addressing data scarcity.
3. **Incorporation of Contextual Information:** Sarcasm often depends on context, tone, and cultural cues, which are difficult to capture through simple sentence-level models. Future research can incorporate more advanced contextual embeddings such as BERT, GPT, or other transformer models to better understand the subtle cues present in the broader context of conversations.
4. **Hybrid Models:** Combining the strengths of various neural network architectures, such as Convolutional Neural Networks (CNN) with RNNs, or experimenting with transformer-based architectures like BERT or GPT, could further enhance the ability to detect sarcasm more effectively. This could also improve the model's ability to detect sarcasm in long or complex sentences.
5. **Real-Time Sarcasm Detection in Dialogue Systems:** Integrating the sarcasm detection model into real-time conversational agents or dialogue systems, such as virtual assistants or chatbots, could improve human-computer interactions by enabling the system to interpret and respond to sarcastic remarks appropriately.
6. **Sentiment-Specific Applications:** Applying the sarcasm detection system to sentiment analysis platforms or content moderation tools could improve the accuracy of sentiment interpretation, particularly for platforms where user-generated content is common, such as social media.
7. **Cross-Domain Adaptability:** Testing the model across different domains, such as news media, entertainment, or education, to understand the adaptability of the sarcasm detection system beyond conversational contexts could open up more practical applications.

By pursuing these directions, future work could significantly improve the effectiveness, scalability, and generalizability of sarcasm detection in low-resource languages.

REFERENCES

[1] Eluri, S., Penmatsa, N.S.L. (2020). Sarcasm detection of sentiments in Telugu language. International Journal of Engineering and Advanced Technology, 10(1): 401-406. https://doi.org/10.35940/ijeat.a1912.1010120

[2] Joshi, A., Bhattacharyya, P., Carman, M. (2017). Automatic sarcasm detection. ACM Computing Surveys, 50(5): 1-22. https://doi.org/10.1145/3124420

[3] Kale, S. (2023). A comprehensive review of sentiment analysis on Indian regional languages: Techniques, challenges, and trends. International Journal on Recent and Innovation Trends in Computing and Communication, 11(9s): 93-110. https://doi.org/10.17762/ijritcc.v11i9s.7401

[4] Marreddy, M., Oota, S., Vakada, L., Chinni, V., Mamidi, R. (2022). Am I a resource-poor language? Data sets, embeddings, models and analysis for four different NLP tasks in Telugu language. ACM Transactions on Asian and Low-Resource Language Information Processing, 22(1): 1-34. https://doi.org/10.1145/3531535

[5] Krishna, M. (2023). Detection of sarcasm using bi-directional RNN based deep learning model in sentiment analysis. Journal of Advanced Research in Applied Sciences and Engineering Technology, 31(2): 352-362. https://doi.org/10.37934/araset.31.2.352362

[6] Oprea, S., Magdy, W. (2019). Exploring author context for detecting intended vs perceived sarcasm. arXiv preprint arXiv:1910.11932. https://doi.org/10.18653/v1/p19-1275

[7] Ganie, A. (2023). Exploring the impact of informal language on sentiment analysis models for social media text using convolutional neural networks. Multidiszciplináris Tudományok, 13(1): 244-254. https://doi.org/10.35925/j.multi.2023.1.17

[8] Soman, S., Swaminathan, P., Anandan, R., Kalaivani, K. (2018). A comparative review of the challenges encountered in sentiment analysis of Indian regional language tweets vs English language tweets. International Journal of Engineering & Technology, 7(2.21): 319. https://doi.org/10.14419/ijet.v7i2.21.12394

[9] Schifanella, R., De Juan, P., Tetreault, J., Cao, L. (2016). Detecting sarcasm in multimodal social platforms. In Proceedings of the 24th ACM international conference on Multimedia, Amsterdam, The Netherlands, pp. 1136-1145. https://doi.org/10.1145/2964284.2964321

[10] Amir, S., Wallace, B.C., Lyu, H., Silva, P.C.M.J. (2016). Modelling context with user embeddings for sarcasm detection in social media. arXiv preprint arXiv:1607.00976. https://doi.org/10.48550/arXiv.1607.00976

[11] Jain, D.K., Kumar, A., Sangwan, S.R. (2022). TANA: The amalgam neural architecture for sarcasm detection in Indian indigenous language combining LSTM and SVM with word-emoji embeddings. Pattern Recognition Letters, 160: 11-18. https://doi.org/10.1016/j.patrec.2022.05.026

[12] Shmueli, B., Ku, L. W., Ray, S. (2020). Reactive supervision: A new method for collecting sarcasm data. arXiv preprint arXiv:2009.13080. https://doi.org/10.18653/v1/2020.emnlp-main.201

[13] Talafha, B., Za'Ter, M.E., Suleiman, S., Al-Ayyoub, M., Al-Kabi, M.N. (2021). Sarcasm detection and quantification in Arabic tweets. In 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), Washington, DC, USA, pp. 1121-1125. https://doi.org/10.48550/arxiv.2108.01425

[14] Zhang, Y., Liu, Y., Li, Q., Tiwari, P., Wang, B., Li, Y., Song, D. (2021). CFN: A complex-valued fuzzy network

for sarcasm detection in conversations. IEEE Transactions on Fuzzy Systems, 29(12): 3696-3710. https://doi.org/10.1109/tfuzz.2021.3072492

[15] Lora, S.K., Shahariar, G.M., Nazmin, T., Rahman, N.N., Rahman, R., Bhuiyan, M. (2022). Ben-SARC: A corpus for sarcasm detection from Bengali social media comments and its baseline evaluation. Preprints, 2: 1-36. https://doi.org/10.31224/osf.io/7yb4c

[16] Aggarwal, A., Wadhawan, A., Chaudhary, A., Maurya, K. (2020). "Did you really mean what you said?": Sarcasm detection in Hindi-English code-mixed data using bilingual word embeddings. In Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020), pp. 7-15. https://doi.org/10.18653/v1/2020.wnut-1.2

[17] Khatri, A. (2020). Sarcasm detection in tweets with BERT and GloVe embeddings. In Proceedings of the Second Workshop on Figurative Language Processing, pp. 56-60. https://doi.org/10.18653/v1/2020.figlang-1.7

[18] Sengupta, A., Suresh, T., Akhtar, M.S., Chakraborty, T. (2022). A comprehensive understanding of code-mixed language semantics using hierarchical transformer. arXiv preprint arXiv:2204.12753. https://doi.org/10.48550/arxiv.2204.12753

[19] Poria, S., Cambria, E., Hazarika, D., Vij, P. (2016). A deeper look into sarcastic tweets using deep convolutional neural networks. arXiv preprint arXiv:1610.08815. https://doi.org/10.48550/arxiv.1610.08815

[20] Ghosh, A., Veale, T. (2016). Fracking sarcasm using neural network. In Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, California, USA, pp. 161-169. https://doi.org/10.18653/v1/w16-0425

[21] Nezhad, Z.B., Deihimi, M.A. (2021). Sarcasm detection in Persian. Journal of Information and Communication Technology, 20(1): 1-20. https://doi.org/10.32890/jict.20.1.2021.6249

[22] Lunando, E., Purwarianti, A. (2013). Indonesian social media sentiment analysis with sarcasm detection. In 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Sanur Bali, Indonesia, pp. 195-198. https://doi.org/10.1109/ICACSIS.2013.6761575

[23] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A. (2010). Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12): 2544-2558. https://doi.org/10.1002/asi.21416

[24] Liu, P., Chen, W., Ou, G., Wang, T., Yang, D., Lei, K. (2014). Sarcasm detection in social media based on imbalanced classification. In Web-Age Information Management: 15th International Conference, WAIM 2014, Macau, China, pp. 459-471. https://doi.org/10.1007/978-3-319-08010-9_49

[25] Desai, N., Dave, A.D. (2016). Sarcasm detection in Hindi sentences using support vector machine. International Journal, 4(7): 8-15.

[26] Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1): 37-46. https://doi.org/10.1177/001316446002000104

[27] Fleiss, J.L., Cohen, J., Everitt, B.S. (1969). Large sample standard errors of kappa and weighted kappa. Psychological Bulletin, 72(5): 323.

[28] Reddy, S., Sharoff, S. (2011). Cross language POS taggers (and other tools) for Indian languages: An experiment with Kannada using Telugu resources. In Proceedings of the Fifth International Workshop on Cross Lingual Information Access, Chiang Mai, Thailand, pp. 11-19.

[29] Bharati, A., Sangal, R., Sharma, D.M., Bai, L. (2006). Anncorra: Annotating corpora guidelines for pos and chunk annotation for Indian languages. LTRC-TR31, 1-38.

[30] Chung, J., Gulcehre, C., Cho, K., Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555. https://doi.org/10.48550/arXiv.1412.3555