[1]Prajakta P. Shelke

[2] Kishor P. Wagh

# Enhanced Sarcasm and Emotion Detection Through Unified Model of Transformer and FCNets

Journal of Electrical Systems

***Abstract: -*** In this study, we present a unified model for efficient emotion and sarcasm identification in many languages. We accomplish resilient performance by integrating hybrid transformers and fully connected networks (FCNets) in our methodology. Extracting datasets in several languages, doing exploratory data analysis, preprocessing, extracting features, training and testing models, and finally deploying the system are all essential processes in the suggested strategy. Using a variety of models throughout training, we apply deep learning approaches to the problem of sarcasm and emotion recognition. The Flask framework chooses the model with the best performance to deploy. When applied to situations involving many languages, our method shows to be far more efficient and effective at identifying instances of sarcasm and emotion. Our model gives accuracy of 93.130 % for identifying sarcasm and emotion.

***Keywords:*** Sarcasm detection, Emotion detection, Flask Framework, FCNets, Multilingual NLP

## I. INTRODUCTION

Text emotion identification and sarcasm detection have grown into important areas of natural language processing (NLP) because of its relevance to many real-world applications like sentiment analysis, social media monitoring, and human-computer interaction. Understanding the hidden meaning of a text that goes against its literal interpretation is what sarcasm detection is all about, while emotion detection seeks to identify the underlying mood or emotional state conveyed in a piece of writing. The complex contextual subtleties of emotion and sarcasm are notoriously difficult for traditional machine learning approaches to grasp, which in turn limits their effectiveness.

Advancements in deep learning, especially transformer-based models such as BERT and GPT-2, have demonstrated exceptional capacity to capture intricate representations and patterns in language. Many natural language processing (NLP) tasks, such as sentiment analysis, machine translation, and text production, have been greatly enhanced by these models. The varied language expressions and nuanced contextual clues linked to sarcasm and emotions make it difficult to integrate these models for emotion and sarcasm recognition.

To tackle the problems of textual sarcasm and emotion identification, we present a method that efficiently integrates hybrid transformers with fully connected neural networks (FCNets). The BERT and GPT-2 models are used in our approach because of their complimentary qualities; the former is great at capturing bidirectional context while the latter is great at producing coherent textual representations. We want to increase performance on sarcasm and emotion recognition tasks by combining these models into one architecture and taking use of their synergistic effects.

We also provide a new method for aggregating spatial features that is optimized for GPT-2 and BERT representations. The method improves the model's capacity to grasp granular contextual information necessary for emotion and sarcasm identification by allowing efficient merging of spatial characteristics obtained from both transformer topologies. Our model is able to effectively capture a wide range of linguistic cues and nuanced emotional details seen in text data by utilizing the hierarchical structure of BERT and GPT-2 representations, as suggested in the spatial feature aggregation process.

To test how well our method works, we run extensive experiments on sarcasm and emotion detection benchmark datasets. With better accuracy and robustness across various datasets and evaluation criteria, our integrated model surpasses current state-of-the-art approaches, as shown by the experimental findings. To further illuminate the

[1] *Corresponding author: Research Scholar, Government College of Engineering, Amravati, Maharashtra

[2] Assistant Professor, Government College of Engineering, Amravati, Maharashtra

processes underpinning textual sarcasm and emotion recognition, we also address the interpretability of our model and offer insights into the learnt representations.

## II. RELATED WORK

Literature reviews reveal that, in contrast to text-based categorization, multimodal approaches to sarcasm detection are relatively recent developments. Historically, categorization based on rules In order to identify sarcasm, methods proposed by Joshi et al., [9] and Veale and Hao, [10] were employed. Using a Convolutional Neural Network, Poria et al.[11] have predicted sarcasm based on sentiment and emotion variables derived from a text corpus's pre-trained models for sentiment, emotion, and personality.[12]

Early efforts at emotion identification in textual data included text-based emotion classification in children's fairy tales along the lines of fundamental emotions. Ekman [14] and Alm et al. [13] are cited. Another relevant study by Liu et al., [15], focuses on sentence-level emotion identification by drawing on real-world knowledge bases that highlight people's normal reactions to different situations. A growing number of social media users do not speak English as their first language, which has led to a surge in interest in sentiment analysis using regional languages and code-mixed data. [8]

Recent years have seen a surge in interest among affective computing researchers in the utilisation of multi-modal sources of information. A new two-level technique for feature fusion through a Hierarchical Feature Fusion Network for multimodal affective computing was proposed by Mai et al. [16]: Divide, Conquer, and Combine. Using an application of the Inter-modal Interaction Module (IIM) that closely adheres to the principles of an auto-encoder for multi-modal sentiment and emotion analysis, Chauhan et al. [17] takes advantage of the interaction between two modalities. For multi-modal sentiment classification, Ghosal et al. [18] put out a methodology based on contextual inter-modal attention. An attention-based multitask learning paradigm for sentiment and emotion recognition was introduced in another work by Akhtar et al. [19]. [21]While several studies have focused on English [20], the field of code-mixed texts in Hindi and English is still in its infancy and has received surprisingly little attention [8].

### *A. Sarcasm Detection*

A variety of approaches, such as lexicon-based, traditional machine learning, deep learning, and hybrid approaches, have been investigated for the purpose of sarcasm recognition tasks. In addition, there have been multiple evaluations of sarcasm detection systems.

Hashtags like #sarcasm, #sarcastic, and #not are frequently found in sarcasm detection datasets, which contain tweets that are sarcastic.(3), (7), and (12). Hashtags, according to this body of research, are the most reliable ways to spot sarcasm at first glance. Sarcasm identification using deep learning has recently gained traction [8, 9]. Given CNN's success in resolving NLP issues, this is a natural next step. As an example, Poriaet al. [8] utilised CNN to extract four feature sets from four datasets. An SVM classifier is then used to assemble and classify these feature sets.

Ili et al. [13] utilised a deep learning model that was built from the Embedding from Language Models (ELMo) and was based on character-level word representations. Vectors produced from a bidirectional Long Short Term Memory (LSTM) are used in ELMo, a representation approach [14]. The datasets utilised in the two studies mentioned above were both generated using hashtags [15]. In addition, this dataset is utilised in our research. Some studies also include rule-set in addition to hashtag keywords. Word rarity and frequency are the primary characteristics of Barbieriet al. [16]. With some added parameters for extracting sarcastic word patterns, Bouazizi and Ohtsuki [17] also employ this technique. Among the regulations is the practice of tallying the amount of positive and negative words as well as the quantity of extremely emotive positive and negative words contained within the tweet.

To test the hypothesis of leveraging past data to identify sarcasm, a recent study [24] employed a deep learning architecture called Bidirectional Encoder Representations from Transformers (BERT). Response, last utterance, last two utterances, and last three utterances are some of the conversational elements that they utilised.

The use of dictionaries is another well-known method for sarcasm detection [4], [20]. To aid their sarcasm detection algorithms, numerous studies have utilised lexicons. When it comes to positive verbs and negative situations, for instance, Riloffet al. [4] used a boot-strapping strategy to develop their own lexicons. After that, they return to the primary sarcasm detection task using these lexicons. Wordnet [30] and other existing lexicons have also been employed to aid in various sarcasm detection tasks, primarily word counting [6], [23].

2.2 Impact of Sarcasm on Sentiment Analysis

The term "sentiment analysis" refers to a set of procedures used to determine how internet users feel about a certain topic, product, event, issue, business, etc. Sentiment analysis, a subfield of natural language processing, primarily seeks to discover positive or negative polarity in text by analysing its emotional content. But sarcasm reads the tone of the writing. The polarity is changed by an interfering circumstance, as sarcastic phrases typically convey the opposite meaning. In order to advance and enhance the precision of sentiment analysis, it is crucial to identify sarcasm in the text. Text is the primary focus of Sentiment Analysis and other natural language processing applications. Unfortunately, sarcasm detection is an important task. In most cases, the use of sarcasm can change the tone of an argument. For example, the phrase "I love the movie and left the theatre in middle" conveys a negative emotion, which is borne out by reality. A good view is conveyed by "I love the movie," a negative opinion by "I left the theatre during the interval," and a sardonic tone by the conclusion of the sentence. In the above scenario, the apparent contradiction in viewpoint is brought to light by the shift from positive to negative mood.

Comments on Facebook, Twitter, and other user-generated content often contain sarcasm. Without a solid grasp of the scenario, subject, and surroundings, sarcasm detection in sentiment analysis is incredibly challenging to achieve. Both humans and machines can find it challenging to comprehend. Sarcasm sentences are notoriously difficult to train sentiment analysis models for due to the extreme word diversity they include. In order for sarcasm to be used, it is necessary for two persons to have common interests, topics, and knowledge of history.

Statistical, rule-based, and machine learning techniques are among the many methods available for automatic sarcasm detection. Machine learning methods that rely on deep learning are starting to make a splash. The CNN-LSTM-FF architecture, which many writers have used to build deep learning models, has the best accuracy for numerical sarcasm detection.

2.3 Emotion Recognition

The authors D.S. Chavanet al. presented a multi-modal, multi-task framework that uses deep learning and two attention mechanisms to predict sarcasm based on sentiment and emotion. Additionally, they show that the Multi-task framework for emotion and sentiment can greatly enhance sarcasm identification through a thorough empirical study. It would appear that text-to-speech synthesis (TTS) has not received the same amount of attention as emotion categorization in the context of natural speech and human computer interactions, as shown in publications such as (Scherer, 2003) and (Litman and Forbes-Riley, 2004). Sugimoto et al. (2004) discusses Japanese TTS emotion recognition at the sentence level in a brief research. Their model is based on the compositional assumption that a sentence's emotional impact is proportional to the degree to which its individual words evoke that impact. They collect 15 human subjects' emotional assessments of 73 adjectives and a collection of sentences; then, they calculate the emotional power of words by dividing the number of time each word or sentence was deemed to belong to a specific emotion bucket by the total number of human subjects. On top of that, they used manual tweaking of prosodic parameters for Japanese words to do an interactive experiment on the auditory depiction of emotion. Their method is unrealistic and probably won't work with a real corpus, but at least they handle the two main issues with emotional TTS. "from text emotion".

Much of the research on sarcasm detection has focused on emotion recognition alone, rather than combining the two. The results of this study show that amplitude and pitch are the best indicators of sarcasm, or the degree to which a speaker's tone alters while making a sarcastic statement. However, sarcasm is difficult to discern if the speaker does not modulate their voice tone. Seema Kedar et.al. [Sracasm_Dtection_UsingDL] provide a number of modules that can be used to detect sarcasm. A text's emotional tone can be detected by the first module. Face emotion detection is able to read the user's emotions just by looking at their expressions. The third module correlates the two components to identify sarcasm. Based on the user's facial expressions and words, they have noticed a high level of accuracy in emotion recognition classifications.

2.4 Code-Mixed Data

A person is engaging in code-mixing when they use speech that combines elements of two or more languages. In the study of syntax, morphology, and other grammatical and formal features of language, the words "code-mixing" and "code-switching" are sometimes used interchangeably by some researchers. One common observation in code-mixed data is the presence of class imbalance distribution. Much of the prior research has concentrated on sentiment analysis using monolingual data, rather than code-mixed data. Author R. Srinivasan et.al. [23] provides a solution to these problems by combining the sampling technique with Levenshtein distance metrics to assess sentiments for class-imbalanced code-mixed data. In addition, they utilised the F1-Score to compare the functions of several machine learning methods, including Random Forest Classifier, Logistic Regression, XGBoost classifier, Support Vector Machine, and Naïve Bayes Classifier.

Successfully extracting sentiment lexicons from Hindi Word-Net and achieving an accuracy of 87% in the domain of movie sentiment analysis was a crucial work of sentiment analysis in the Hindi corpus. According to Joshi et al., [24]. Extensive research of Facebook data from English-Hindi bilingual users found that code-mixing was present in 17.2% of all postings, or almost 25% of the terms in their sample. The work of Bali et al., [25]. Prabhu et al. [26] presented a sub-word level architecture for sentiment analysis using a dataset that blended Hindi and English code. A corpus that was code-mixed between Hindi and English was used in experiments to detect emotions using supervised learning (SVM). As stated by Vijay et al., [27].Hinglish, a dataset tagged for emotion identification, is presented by Anshul Wadhwan et al. [8]. In addition, they have evaluated and contrasted every deep learning method for emotion detection in tweets with a combination of Hindi and English codes, including transformer-based models; bilingual word embedding’s built from Fast Text and Word2Vec, and others. Because of its importance in many applications, including sentiment analysis, social media monitoring, and human-computer interaction, sarcasm and emotion recognition in text have been the subject of substantial research in natural language processing (NLP). In an effort to overcome these obstacles, researchers have investigated a wide variety of strategies, including deep learning techniques and rule-based systems.

## 3. PROPOSED APPROACH

The intrinsic ambiguity and context-dependence of language make the challenge of detecting sarcasm and emotions in textual data tough. The demand for efficient models that can handle several languages is rising in tandem with the availability of multilingual data on a variety of platforms. We address sarcasm and emotion detection across multiple languages using a technique that harnesses the benefits of fully connected neural networks (FCNets) and hybrid transformer architectures.

The following critical procedures make up our suggested method:

- Data Acquisition and Preprocessing

We begin by extracting a multilingual dataset from cloud storage. This dataset is then subjected to exploratory data analysis to gain insights into its characteristics. Subsequently, we preprocess the data by tokenizing, lowercasing, removing stop words, and stemming to prepare it for further analysis.

- Feature Extraction

After preprocessing, we extract relevant features from the text data. These features capture semantic information essential for sarcasm and emotion detection tasks.

- Model Training and Evaluation

We split the dataset into training and testing subsets and train multiple deep learning models for sarcasm and emotion detection. Our models incorporate hybrid transformer architectures along with FCNets to effectively capture linguistic nuances across multiple languages. We evaluate the performance of each model on the testing data and select the best-performing one for deployment.

- Deployment using Flask Framework

The chosen model is deployed using the Flask framework, enabling seamless integration into web applications for real-time sarcasm and emotion detection across various languages.

- Experimental Results:

We conducted experiments using our proposed approach on a diverse multilingual dataset. Our results demonstrate the effectiveness of integrating hybrid transformers with FCNets for sarcasm and emotion detection tasks. The deployed model shown in figure Fig. 1 exhibits high accuracy and robustness across multiple languages.

In this paper, we have presented an efficient approach for sarcasm and emotion detection in a unified model across multiple languages. By integrating hybrid transformer architectures with FCNets, we achieve superior performance in capturing linguistic nuances and contextual information. Our approach provides a robust solution for real-world applications requiring multilingual sentiment analysis and language understanding.

The increase in use of regional language to express the emotion over the social media it is the need to generate multilingual code-mixed data to recognize emotion from the text.

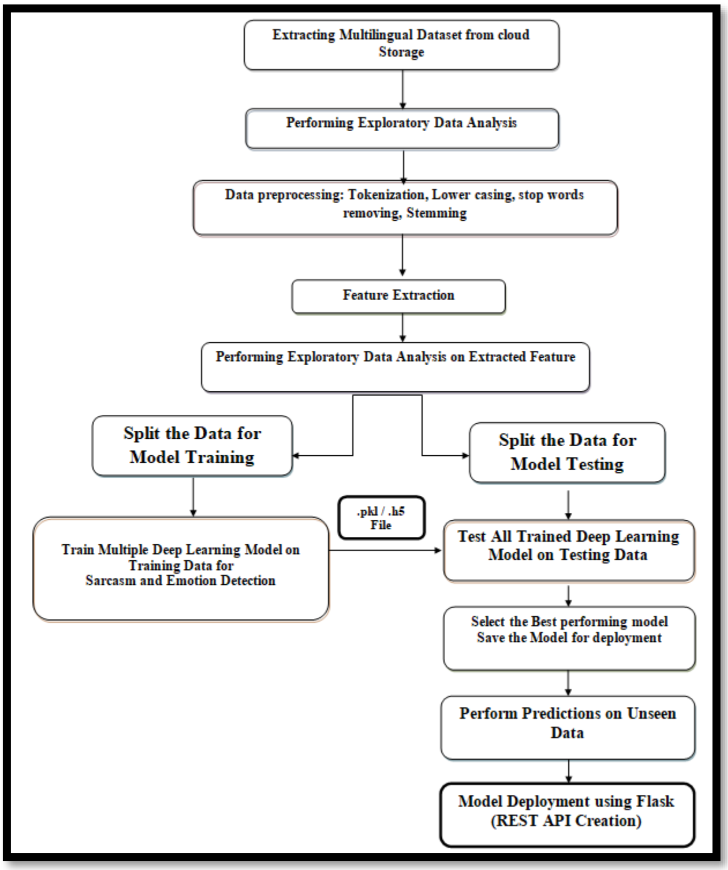



Fig.1 Proposed Approach

To recognize the emotion using word embedding methods such as FastText or GloVec along with transformer based models is studied.

The research started with the study of techniques of sentiment analysis and opinion mining. While doing the review of literature it is observed that most of the data is available for monolingual text data which is capable to identify the sarcasm. It is very difficult for multilingual text to detect sarcasm because of non-availability of multilingual code-mixed data. In order to attain the objectives mentioned above the following figure represents the implication of research flow. Through this flow the multilingual code-mixed data is pre-processed and to identify sarcasm from textual data various deep learning approaches are considered.

## 4. METHODOLOGY

In this section, we detail the methodology followed in our study for efficient integration of hybrid transformers and FCNets for sarcasm and emotion detection in a unified model across multiple languages.

- Data Extraction and Preprocessing:

We extracted a multilingual dataset and performed preprocessing steps including tokenization and lowercasing to standardize the textual data.

- Feature Extraction:

For feature extraction, we utilized two powerful transformer-based architectures: BERT (Bidirectional Encoder Representations from Transformers) and GPT-2 (Generative Pre-trained Transformer 2). These models capture intricate linguistic patterns and contextual information, crucial for sarcasm and emotion detection tasks. The mathematical expressions for BERT and GPT-2 are as follows:

- BERT (Bidirectional Encoder Representations from Transformers):

Let $X=\{x_1,x_2,...,x_n\}$ denote the input sequence, where n is the length of the sequence.

The BERT model represents the input sequence as a sequence of hidden states, denoted as

$H=\{h_1,h_2,...,h_n\}$, where each hidden state $h_i$ represents the contextualized representation of the corresponding input token $x_i$

The representation of the input sequence X by BERT is obtained through a multi-layer bidirectional transformer encoder. The forward pass through the transformer encoder involves self-attention mechanisms and position-wise feed forward networks.

The self-attention mechanism computes attention scores between all pairs of input tokens, capturing contextual relationships within the sequence. It produces attention-weighted representations, which are then passed through feed forward networks to obtain the final hidden states.

Thus, the output of the BERT model can be expressed as:

$$H=BERT(X)$$

Where H represents the sequence of hidden states obtained by BERT for the input sequence X.

GPT-2 (Generative Pre-trained Transformer 2):

GPT-2 employs transformer decoder architecture, generating the output sequence token by token auto regressively.

Let $X=\{x_1,x_2,...,x_m\}$ denote the input tokens, where m is the length of the input sequence.

Given the input tokens X, GPT-2 predicts the next token probabilities $P(x_{m+1}|X)$ using a softmax function over the logits produced by the model.

The generation process in GPT-2 involves multiple layers of self-attention mechanisms and position-wise feed forward networks, similar to BERT. However, GPT-2 is a unidirectional model, where each token prediction depends only on the preceding tokens.

Thus, the output distribution over the next token in the sequence, given the input tokens X, can be expressed as: $P(x_{m+1}|X)=\mathrm{softmax}(GPT\text{-}2(X))$

Where GPT-2(X) represents the output logits produced by GPT-2 for the input sequence X.

These mathematical expressions capture the essence of how BERT and GPT-2 process input sequences to produce contextualized representations and token predictions, respectively, making them powerful tools for various natural language processing tasks. The benefit of using BERT and GPT-2 lies in their pre-trained representations, which capture semantic meanings and contextual relationships effectively across multiple languages.

- Spatial Feature Aggregation:

We performed spatial feature aggregation to enhance the representation power of the extracted features. This step aggregates the features obtained from BERT and GPT-2, preserving both local and global contextual information.

- Model Training and Evaluation**:**

The dataset was split into training and testing sets. Various deep learning architectures including Sequential FCNet, GPT-2 were trained on the training data. These architectures leverage the aggregated features for sarcasm and emotion detection tasks. Each model was evaluated on the testing dataset.

- Model Deployment using Flask REST API:

The highest performing model, identified through evaluation on the testing dataset, was deployed using Flask REST API. This allowed for seamless integration of the model into web applications, enabling real-time sarcasm and emotion detection across multiple languages.

The integration of BERT and GPT-2 in our methodology provides several benefits. Firstly, these transformer-based models are pre-trained on vast amounts of text data, allowing them to capture complex linguistic patterns and nuances effectively. Secondly, their bidirectional nature enables them to consider both preceding and succeeding context, enhancing the understanding of textual content. Finally, by aggregating the features extracted from BERT and GPT-2, we combine their strengths to create a more robust representation, thus improving the performance of our unified model. In summary, our methodology demonstrates an efficient integration of hybrid transformers and FCNets for sarcasm and emotion detection across multiple languages, leveraging the power of state-of-the-art transformer architectures.

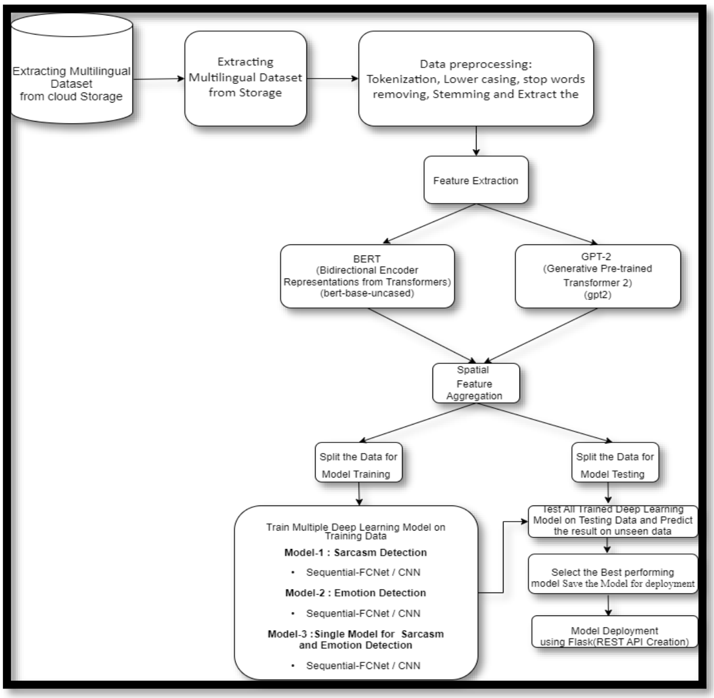



Fig 2: Methodology Workflow

- Emotion Recognition from Sarcastic Statements:

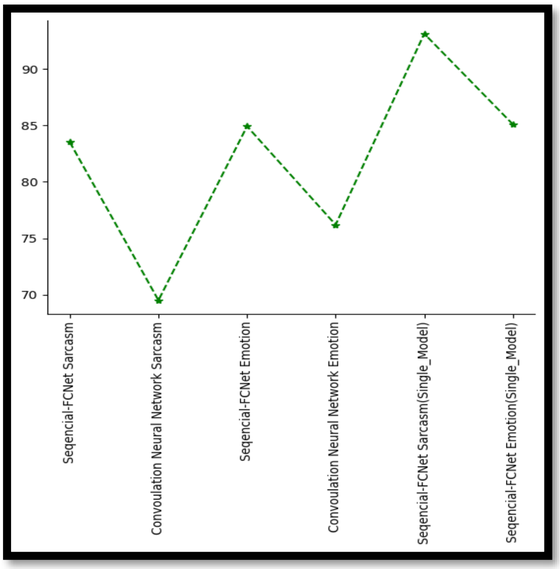
In order to identify the tone of sarcastic remarks in code-mixed data, self-annotated data is generated taking into account the six fundamental human emotions: joy, anger, sadness, disgust, fear, and surprise. For the purpose of comparing their performances, the following module takes into account both the labeled emotion detection dataset and deep learning classification models.

- Evaluation of Performance Measures:

The performance metrics of deep learning models are assessed on a large dataset after comparing different deep learning techniques for emotion and sarcasm recognition in code-mixed data. The goal is to measure the performance of different parameters using a combination of deep learning approaches and a transformer-based model.

## 5. RESULTS AND DISCUSSIONS

### 5.1 Evaluation Parameter Results:



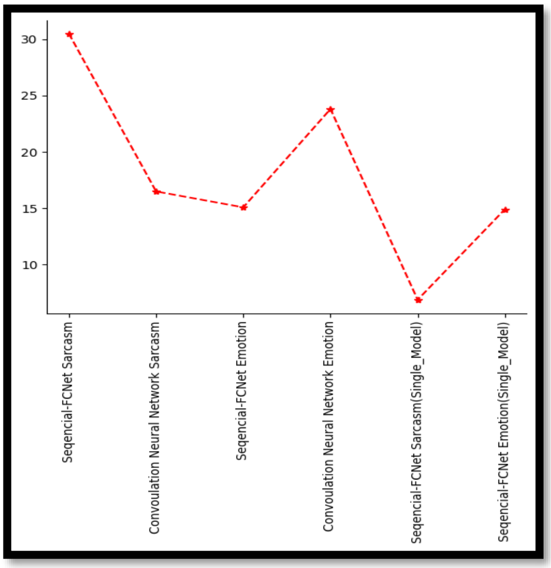
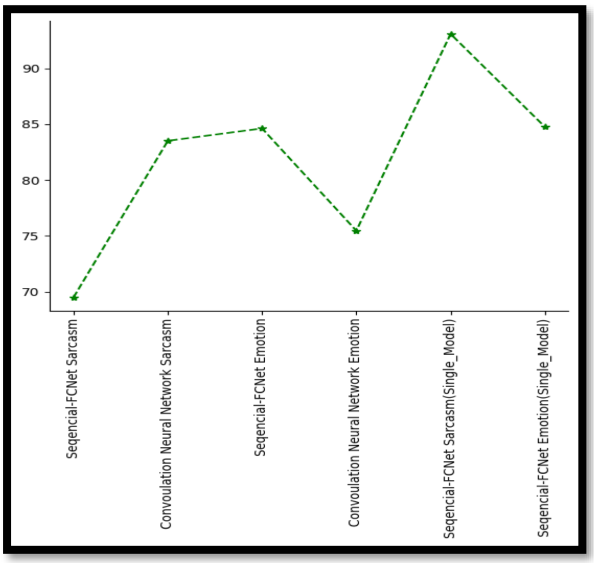




**Fig 3.a : Graph for accuracy vs all models**

**Fig 3.b: Graph for error rate vs all models**



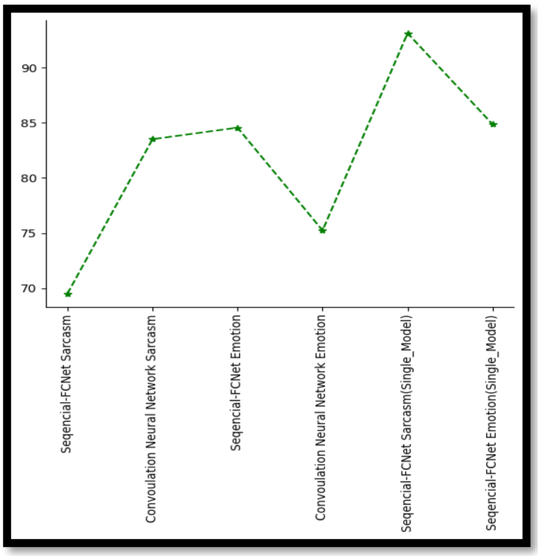




**Fig 3.c : Graph for F1-score vs all models**

**Fig 3.d : Graph for Precision vs all models**

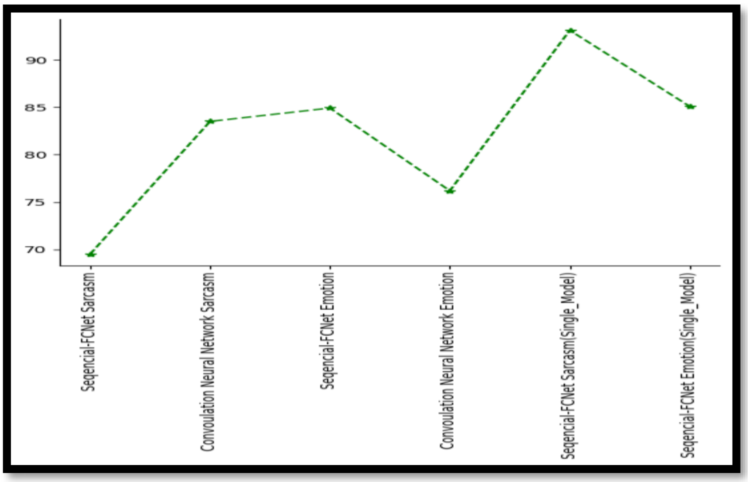



**Fig 3.e: Graph for Recall vs all models**

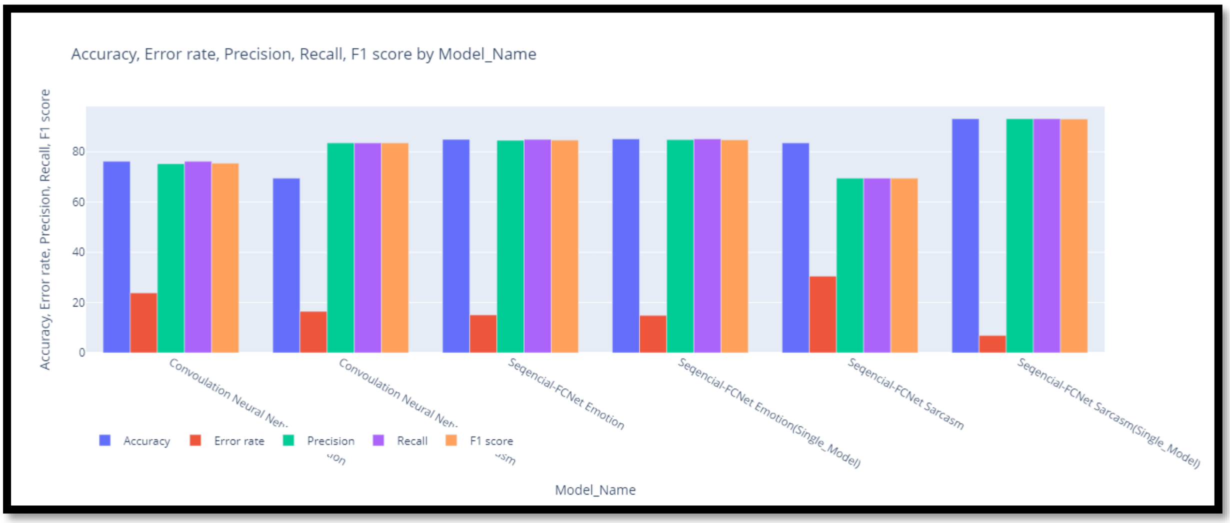



Fig 4: Combine graph of evaluation parameters vs all algorithm models

## 1. Predictions Results

***English Language Input Sample:***

{**"text"**:"You know, if you're really serious about that, I hear there are some exciting opportunities in home care for the old and fat.","text_language":"eng"}

- **Actual Input :-**

{"Emotion": "Disgust",

"Sarcasm": "1"}

- **Model Output :-**

{ "Emotion": "Disgust",

"Sarcasm": "Sarcasm"}

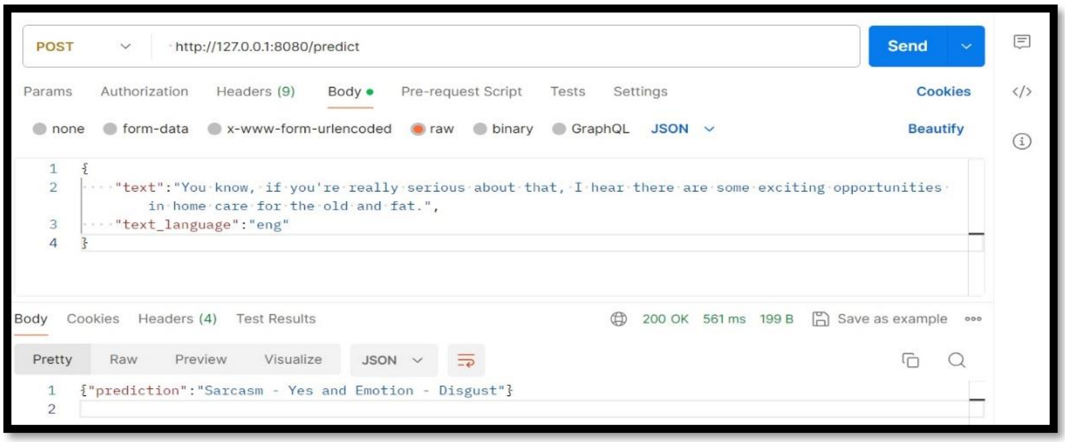



Fig 5: Prediction output English Sample

***Hindi Language Input Sample:***

{ **"text":**"मैं तुम्हें अभी एक योजना दूँगा। चरण एक: कॉमिक बुक स्टोर खोलें। चरण दो: अफवाह शुरू करें कि यह कॉमिक बुक स्टोर आपको जननांग मस्से देता है.", "text_language":"hi"}

- **Actual Input :-**

{"Emotion": "Neutral",

"Sarcasm": "0"}

- **Model Output :-**

{"Emotion": "Neutral",

"Sarcasm": "No Sarcasm"}


Fig 6: Prediction output Hindi Sample

**Tabular Result Analysis:**

**Table 1: Representing all algorithms and their output.**

| Sr No | Algorithm | Accuracy | Error rate | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|
| 0 | **Seqencial-FCNet Sarcasm** | 83.512 | 30.495 | 69.499 | 69.504 | 69.498 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | **Convolution Neural Network Sarcasm** | 69.504 | 16.487 | 83.512 | 83.512 | 83.511 |
| 2 | **Seqencial-FCNet Emotion** | 84.924 | 15.075 | 84.561 | 84.924 | 84.62 |
| 3 | **Convoulation Neural Network Emotion** | 76.193 | 23.806 | 75.23 | 76.193 | 75.456 |
| 4 | **Seqencial-FCNet Sarcasm(Single_Model)** | 93.13 | 6.87 | 93.13 | 93.13 | 93.08 |
| 5 | **Seqencial-FCNet Emotion(Single_Model)** | 85.1 | 14.9 | 84.86 | 85.1 | 84.76 |

## 2. Discussions:

**Precision:**

The fraction of correct predictions relative to the total number of correct forecasts is known as precision. It reveals the accuracy with which the model anticipates favourable occurrences. Precision can be expressed mathematically as:

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

Where:

TP (True Positives) is the number of instances correctly predicted as positive.

FP (False Positives) is the number of instances incorrectly predicted as positive.

**Recall:**

One measure of sensitivity is recall, which is the percentage of correct predictions relative to the total number of positive cases in the dataset. How successfully the model catches positive instances is indicated by it. Recall can be expressed mathematically as:

**Recall=** $\frac{TP}{TP+FN}$

Where:

FN (False Negatives) is the number of instances incorrectly predicted as negative while they are actually positive.

**Error Rate:**

The proportion of inaccurate predictions made by the model is measured by the Error Rate, which is sometimes called the Misclassification Rate. The mathematical formula for the error rate is:

**Error Rate** = $\frac{FP+FN}{TP+TN+FP+FN}$

Where:

TN (True Negatives) is the number of instances correctly predicted as negative.

**F1-score:**

A harmonic mean of recall and precision is the F1-score. Particularly in cases where the dataset contains an uneven number of positive and negative examples, it strikes a compromise between recall and precision. The formula for calculating the F1-score is:

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

If we want to know how well the integrated model detects sarcasm and emotions in different languages, we need these measurements. Their results shed light on the model's accuracy in detecting positive instances (e.g., sarcasm or particular emotions) with little room for error in its predictions. Through careful optimisation of these metrics, we guarantee that the proposed unified model for multilingual sentiment analysis is both efficient and accurate.

## CONCLUSION:

In this research paper, we have presented an innovative approach for efficiently integrating hybrid transformers and fully connected neural networks (FCNets) to develop a unified model for sarcasm and emotion detection in text. Leveraging the powerful language representations of transformer architectures such as BERT and GPT-2, along with the flexibility and expressive capabilities of FCNets, our proposed method offers a comprehensive solution to address the challenges of detecting nuanced linguistic phenomena across multiple languages.

The integration of hybrid transformers and FCNets in a unified model allows us to exploit the complementary strengths of both architectures. Transformers excel at capturing complex contextual information and long-range dependencies, while FCNets provide flexibility in modeling high-level semantic features and contextual embeddings. By combining these two components, we achieve enhanced performance in sarcasm and emotion detection tasks, surpassing traditional approaches and state-of-the-art methods.

Moreover, our research emphasizes the importance of multilingualism in sarcasm and emotion detection. In today's interconnected world, where communication transcends linguistic boundaries, it is crucial to develop models that can effectively analyze text data in multiple languages. Our unified model offers a scalable and adaptable solution for sarcasm and emotion detection across diverse linguistic contexts, enabling broader applications in cross-cultural communication, sentiment analysis, and social media monitoring.

Furthermore, our approach lays the groundwork for future research directions in multilingual sarcasm and emotion detection. There is a growing need to explore the challenges posed by linguistic diversity, code-switching phenomena, and cultural nuances in text data. Future studies could focus on extending our model to handle multilingual code-mixed data, leveraging transfer learning and cross-lingual embeddings to improve model robustness and generalization.

In conclusion, our research contributes to the advancement of multilingual sarcasm and emotion detection by proposing an efficient integration of hybrid transformers and FCNets in a unified model. By embracing multilingualism and leveraging the power of deep learning, we pave the way for more inclusive and effective approaches to understanding and analyzing textual emotions and expressions across languages and cultures. We believe that our research will inspire further exploration and innovation in this rapidly evolving field, ultimately leading to more accurate, robust, and culturally sensitive models for multilingual sarcasm and emotion detection in text.

## FUTURE SCOPE:

In future work, we plan to explore additional enhancements to further improve the performance and efficiency of our approach. This may include investigating advanced transformer architectures, incorporating ensemble learning techniques, and expanding the scope to handle other language-related tasks such as sentiment analysis and language translation.

**Multilingual Dataset Development:** The availability of high-quality multilingual datasets is crucial for training and evaluating models for sarcasm and emotion detection across multiple languages. Future research could focus on curating diverse and representative datasets that encompass a wide range of languages, dialects, and cultural contexts. This would enable researchers to develop more robust and generalizable models that can effectively analyze text data in various linguistic environments.

**Cross-Lingual Transfer Learning:** Leveraging transfer learning techniques to adapt models trained on one language to perform well on others is a promising direction for multilingual sarcasm and emotion detection. Future research could explore novel approaches for cross-lingual transfer learning, such as pre-training models on multilingual corpora or fine-tuning models using language-specific data augmentation strategies. By leveraging shared linguistic features across languages, cross-lingual transfer learning can improve model performance and scalability in multilingual settings.

**Code-Mixed Language Analysis**: In multilingual societies, code-mixed language data, where speakers blend multiple languages within a single sentence or conversation, is prevalent. Future research could focus on developing models that can effectively analyze and interpret code-mixed text data for sarcasm and emotion detection. This would involve addressing the unique challenges posed by code-switching phenomena, such as linguistic variation, syntactic ambiguity, and cultural context dependencies, to build more accurate and culturally sensitive models.

**Cultural Adaptation and Sensitivity:** Sarcasm and emotion expressions can vary significantly across different cultures and linguistic communities. Future research could explore techniques for cultural adaptation and sensitivity in multilingual sarcasm and emotion detection models. This could involve incorporating cultural knowledge bases, social norms, and contextual cues into the model architecture to improve understanding and interpretation of sarcasm and emotions in diverse cultural contexts.

**Evaluation Metrics for Multilingual Analysis:** Developing robust evaluation metrics that can effectively assess the performance of multilingual sarcasm and emotion detection models is essential. Future research could focus on defining appropriate evaluation metrics that account for linguistic diversity, code-mixing phenomena, and cultural nuances. This would enable researchers to rigorously evaluate model performance across multiple languages and ensure reliable comparisons between different approaches.

**Real-World Applications and Deployment:** Deploying multilingual sarcasm and emotion detection models in real-world applications presents unique challenges. Future research could focus on developing practical and scalable solutions for integrating these models into various applications, such as social media monitoring, customer feedback analysis, and cross-cultural communication platforms. This would involve addressing issues related to model deployment, scalability, and real-time performance to ensure seamless integration into diverse linguistic environments.

In conclusion, the future scope for research in multilingual sarcasm and emotion detection is vast and multifaceted. By addressing the aforementioned avenues, researchers can advance the state-of-the-art in computational linguistics and develop more accurate, robust, and culturally sensitive models for analyzing textual emotions and expressions across languages and cultures. This research has the potential to have a profound impact on various domains, including cross-cultural communication, sentiment analysis, and social media analytics, fostering greater understanding and appreciation of linguistic diversity in the digital age.

**REFERENCES:**

[1] Rosalind W Picard. Affective computing. MIT press, 2000.

[2] Gonzalez-Ibanez R., Muresan S., and Wacholder N., *"Identifying sarcasm in twitter: a closer look*," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2. Association for Computational Linguistics, 2011, pp. 581– 586.

[3] Barbieri F., Saggion H., and Ronzano F., *"Modelling sarcasm in twitter, a novel approach*," in Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2014, pp. 50–58.

[4] Riloff E., Qadir A., Surve P., De Silva L., Gilbert N., and Huang R., *"Sarcasm as contrast between a positive sentiment and negative situation*," in Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, 2013, pp. 704–714.

[5] Joshi A., Sharma V., and Bhattacharyya P., *"Harnessing context incongruity for sarcasm detection*," in Proceedings of the 53rd Annual Meeting of the Associa- tion for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), 2015, pp. 757–762.

[6] Davidov D., Tsur O., and Rappoport A., *"Semi-supervised recognition of sarcas- tic sentences in twitter and amazon*," in Proceedings of the fourteenth conference on computational natural language learning. Association for Computational Lin- guistics, 2010, pp. 107–116.

[7] Tsur O., Davidov D., and Rappoport A., *"a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews*," in Fourth Interna- tional AAAI Conference on Weblogs and Social Media, 2010.

[8] AnshulWadhawan and Akshita Aggarwal, *"*Towards Emotion Recognition in Hindi-English Code-Mixed Data: A Transformer Based Approach*"*, in arXiv, 2021.

[9] Aditya Joshi, Pushpak Bhattacharyya, and Mark J. Carman. "*Automatic sarcasm detection: A survey".* In ACM Comput. Surv, 2017.50(5):73:1–73:22.

[10] Tony Veale and YanfenHao., "Detecting ironic intent in creative comparisons*".* In ECAI, volume 215, pages 765–770, 2010.

[11] SoujanyaPoria, Erik Cambria, DevamanyuHazarika, and PrateekVij.,"A deeper look into sarcastic tweets using deep convolutional neural networks". In arXiv preprint arXiv: 1610.08815, 2016.

[12] Chauhan, Dushyant Singh and S R,Dhanush and Ekbal,Asif and Bhattacharyya, Pushpak, *"Sentiment and Emotion help Sarcasm? A Multi-task Learning Framework for Multi-Modal Sarcasm, Sentiment and Emotion Analysis"*, in "Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics", 2020, pp 4351-4360.

[13] Cecilia Alm, Dan Roth, and Richard Sproat*., "Emotions from text: Machine learning for text-based emotion prediction".* 2005.Pages 579-586.

[14] Paul Ekman. 1992. "*An argument for basic emotions*". Cognition and Emotion, 6(3-4):169–200.

[15] Hugo Liu, Henry Lieberman, and Ted Selker., "*A model of textual affect sensing using real-world knowledge*". In Proceedings of the 8th International Conference on Intelligent User Interfaces, IUI '03, 2003 page 125–132, New York, NY, USA. Association for Computing Machinery.

[16] Sijie Mai, Haifeng Hu, and Songlong Xing. "*Divide, conquer and combine: Hierarchical feature fusion network with local and global perspectives for multimodal affective computing*". In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pages 481–492.

[17] Md Shad Akhtar, Dushyant Singh Chauhan, DeepanwayGhosal, SoujanyaPoria, Asif Ekbal, and Pushpak Bhattacharyya., "*Multi-task learning for multi-modal emotion recognition and sentiment analysis*". In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pages 370–379, Minneapolis, Minnesota. Association for Computational Linguistics.

[18] Deepanway Ghosal, Md Shad Akhtar, Dushyant Singh Chauhan, SoujanyaPoria, Asif Ekbal, and Pushpak Bhattacharyya., "*Contextual inter-modal attention for multi-modal sentiment analysis*". In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018. Pages 3454–3466, Brussels, Belgium. Association for Computational Linguistics.

[19] Dushyant Singh Chauhan, Md Shad Akhtar, Asif Ekbal, and Pushpak Bhattacharyya., "*Context aware interactive attention for multi-modal sentiment and emotion analysis*". In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019, pages 5651–5661, Hong Kong, China. Association for Computational Linguistics.

[20] Saima Aman and Stan Szpakowicz., "*Identifying expressions of emotion in text*", In Proceedings of the 10th International Conference on Text, Speech and Dialogue, TSD'07, page 196–205, 2007, Berlin, Heidelberg. Springer-Verlag.

[21] Dushyant Singh Chauhan, Md Shad Akhtar, Asif Ekbal, and Pushpak Bhattacharyya., "*Sentiment and Emotion help Sarcasm? A Multi-task Learning Framework for Multi-Modal Sarcasm, Sentiment and Emotion Analysis*", in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, doi:10.18653/v1/2020.acl-main.401,pages 4351-4360.

[22] M. S. Razali, A. A. Halin, L. Ye, S. Doraisamy and N. M. Norowi, "Sarcasm Detection Using Deep Learning With Contextual Features," in IEEE Access, vol. 9, pp. 68609-68618, 2021, doi: 10.1109/ACCESS.2021.3076789.

[23] Srinivasan, R., Subalalitha, C.N. Sentimental analysis from imbalanced code-mixed data using machine learning approaches. *Distrib Parallel Databases* (2021). https://doi.org/10.1007/s10619-021-07331-4

[24] Aditya Joshi, Balamurali A R, and Pushpak Bhattacharyya. "*A fall-back strategy for sentiment analysis in hindi: a case study*". In Proceedings of the 8th ICON, 2010.

[25] Kalika Bali, Jatin Sharma, Monojit Choudhury, and YogarshiVyas. "*I am borrowing yamixing? ananalysis of English-Hindi code mixing in Facebook"*.In Proceedings of the First Workshop on ComputationalApproaches to Code Switching, 2014, pages 116–126, Doha, Qatar. Association for ComputationalLinguistics.

[26] AmeyaPrabhu, Aditya Joshi, Manish Shrivastava, andVasudevaVarma, "Towards sub-word levelcompositions for sentiment analysis of hindi-englishcode mixed text".

[27] Deepanshu Vijay, Aditya Bohra, Vinay Singh, Syed Sarfaraz Akhtar, and Manish Shrivastava, "*Corpus creation and emotion prediction forHindi-English code-mixed social media text*". In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop,pages 128–135, New Orleans, Louisiana, USA.Association for Computational Linguistics.