International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Comparative Analysis of Fine-Tuned BERT and Fine-Tuned RoBERTa for Sarcasm Detection on Social Media

Raksha Kushwaha[1] , Mansi Sharma[2] , Mr. Amit Kumar Srivastava[3]

[1,2]Student, Computer Science, National P.G. College, Lucknow, Uttar Pradesh, India

[3] Assistant Professor, Computer Science, National P.G. College, Lucknow, Uttar Pradesh, India

kushwaharaksha1028@gmail.com sharmamansi0009@gmail.com amit_sri_in@yahoo.com

ABSTRACT :

Sarcasm is a language term with the meaning opposite to what is actually being said, thereby achieving the intention of ridicule, taunt, or conveyance of contempt. Sarcasm is a complex social phenomenon that is identifiable by its tone of voice, exaggeration, or context. Because sarcastic discourse is subtle, sarcasm detection in NLP is a substantial challenge. This paper presents a comparative study on the performance of two pre-trained transformer-based models, BERT (Bidirectional Encoder Representations from Transformers) and RoBERTa (A Robustly Optimized BERT Pretraining Approach), when fine-tuned to carry out sarcasm detection tasks on a large-scale labelled dataset. We compare both models using evaluation metrics like their accuracy, precision, recall, and F1-score. The findings from the experiment show that while BERT achieves a high overall accuracy of 93%, RoBERTa underperforms significantly with an accuracy of 56%, mainly due to its inability to detect sarcastic instances particularly when dealing with complicated datasets that require contextual knowledge. This research emphasizes the significance of model selection in order to fine-tune pre-trained transformer-based models for sarcasm NLP tasks.

Keywords: Sarcasm detection, NLP, BERT Model, RoBERTa Model, Fine-tuning.

Introduction :

Sarcasm, sometimes referred to as a form of verbal irony, plays a crucial role in human communication [1] as it conveys quite the opposite meaning of what it actually means. It is an common part of everyday interactions and is quite often used to express amusement, mockery, or irritation. Despite its ubiquity, sarcasm detection still creates a great challenge for natural language processing tasks, be they sentiment-based or emotion-based. Detecting sarcasm is an important aspect when it comes to analysing sentiments and opinion mining especially for social media platforms as users tend to express their sentiments in an indirect manner. The detection of sarcasm is necessary as sarcastic comments

can tend to obscure the meaning of the text resulting in wrong sentiment analysis as it allows individuals to convey contempt and negativity under the guise of overt positive representation, making it crucial to discern true sentiments and beliefs [2] .

Traditional approaches often struggle to capture the nuanced interplay between the text, context, and underlying intent that characterizes sarcastic expressions [3]. Recent advances in pre-trained transformer models, such as BERT (Bidirectional Encoder Representations from Transformers) and RoBERTa (Robustly Optimized BERT Pretraining Approach), have achieved state-of-the-art results in a variety of NLP tasks. However, the effectiveness of these models for sarcasm detection, specifically when fine-tuned, has not been extensively studied.

This paper aims to compare the performance of BERT and RoBERTa when fine-tuned on sarcasm detection, analysing their performance through various evaluation metrics and confusion matrices. The key contributions include a comparative analysis of BERT and RoBERTa-based models for sarcasm detection, providing insights into their respective strengths and limitations.

Related Work :

The detection of sarcasm has garnered significant attention in Natural Language Processing (NLP). Initially, research developed rule-based systems, lexicons, and supervised learning approaches of SVMs and logistic regression. With the birth of deep learning, recurrent neural networks, convolution neural networks, and most recently transformers have taken sarcasm detection to a new level.

The challenges of sarcasm detection have been addressed in a new way with the advent of pre-trained transformer-based architectural language models such as BERT and RoBERTa for enhanced performance. These models have demonstrated state-of-the-art performance in several NLP tasks, and they are especially well-suited for sarcasm detection due to their contextual awareness.

BERT, introduced by Devlin et al. [4], is one of the most influential transformer models that pre-trains a language model bidirectionally. For instance, recent studies have demonstrated that models like BERT and its variants (DistilBERT, BERT-large) outperform traditional machine learning methods in sarcasm detection tasks by achieving higher precision and recall scores [8].

Furthermore, hybrid models combining BERT with other deep learning techniques, such as Long Short-Term Memory (LSTM) networks, have also achieved significant improvements in detecting sarcasm in short texts, such as tweets and news headlines [9]. Hybrid models, such as Opinion-BERT, leverage BERT's embeddings along with other neural layers, like CNNs and BiGRUs, to improve performance in multi-task scenarios, particularly for understanding sentiments and user statuses from social media text [12](Hossain et al., 2025).

On the other hand, RoBERTa, a robustly optimized variant of BERT, introduced by Liu et al. [5], improves BERT's performance through modifications in pre-training tasks, data sizes, and hyperparameters. Furthermore, RoBERTa benefits from dynamic masking, which means that during the training process, the masking patterns change, allowing for better generalization and also it is trained with much larger dataset and for longer durations, allowing the model to grasp language nuances better.

All these changes contribute in improving RoBERTa's state-of-art performance across various NLP tasks. For instance, RoBERTa achieved 98.46% accuracy in sentiment polarity analysis when applied to public opinion comments [13](Wang et al., 2025). Additionally, RoBERTa has been adapted for specialized domains, such as legal contexts, where the model has been fine-tuned on domain-specific corpora like the "LegalPT" dataset, resulting in RoBERTaLexPT, a model that outperforms general models in Portuguese legal text analysis [14](Garcia et al., 2024).

Its enhanced training methodology, which involves dynamic masking and training with larger datasets, has led to more accurate sarcasm detection when compared to the original BERT model [10]. Additionally, RoBERTa has been successfully integrated with multi-modal sarcasm detection systems, utilizing text and images for richer context understanding [11].

In the context of sarcasm detection, Dadu and Pant [6] leveraged RoBERTa with context separators, showing promising results in online discourse analysis. Comparative studies, such as Mao and Liu [7] in the related field of sarcasm detection, have highlighted the potential advantages of RoBERTa over BERT in capturing nuanced linguistic features. Both models have been applied to various NLP tasks, but there has been limited exploration of their comparative performance for sarcasm detection.

Despite these advancements, several challenges remain in sarcasm detection, particularly in developing models that can effectively generalize across different contexts. Built upon these foundations, this paper particularly focuses on addressing the challenges of context-sensitivity and model optimization in sarcasm detection. It leverages the strengths of both BERT and RoBERTa as base models, incorporating a sophisticated feature extractor and classifier architecture. The approach implements fine-tuning technique to optimize model performance.

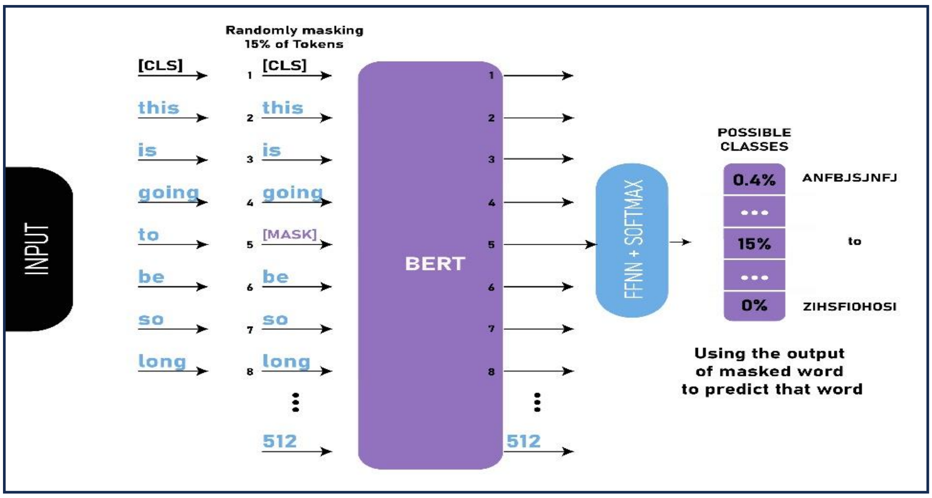
Bert and RoBERTa : Overview :

a) Bert History and Architecture:

BERT (Bidirectional Encoder Representations from Transformers), introduced by Devlin et al. (2019)[4], is a breakthrough in language modelling by applying transformers [15]. BERT is designed to pre-train deep bidirectional representations from unlabelled text by jointly conditioning on both left and right context in all layers[4],. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications[4]. The components of BERT's architecture are:

- A **stack of transformer encoders**, where each layer performs self-attention and feed-forward transformations.
- **Self-attention mechanisms** that assign varying importance to words based on their context, allowing BERT to weigh the relationship between words bidirectionally.
- The model is pre-trained using two unsupervised objectives: **Masked Language Modeling (MLM)**, where some words in a sentence are masked, and the model must predict them, and **Next Sentence Prediction (NSP)**, where the model determines if two sentences are sequential in a document [16].

Fig. 1 Diagrammatic representation of BERT architecture [8].



The BERT base model consists of 12 transformer layers, each with 12 attention heads and 768 hidden dimensions, while the larger BERT-large model has 24 layers, 16 heads, and 1024 hidden dimensions. This architecture allows BERT to model complex syntactic and semantic relationships in text.

Fine-Tuning BERT

Fine-tuning BERT refers to fine-tuning the pre-trained BERT model for a specific downstream task [8], for example, sentiment analysis, named entity recognition (NER), or question answering, through further training on task-specific, relatively small dataset.

In practice, the process of fine-tuning involves:

1. **Loading the pre-trained BERT model**: The base BERT model (pre-trained on general text) is loaded.
2. **Adding task-specific layers**: For different tasks, specific layers (e.g., a classification head) are added on top of BERT's architecture.
3. **Training on the new dataset**: The model is further trained on a labelled dataset specific to the new task, adjusting the model weights to better suit the task.
4. **Optimization**: Hyperparameters like learning rate, batch size, and epochs are fine-tuned to ensure the model performs well on the specific task.

During fine-tuning, the entire BERT model (or even individual layers) gets updated through further training, allowing it to learn the task specifics but retaining its base language understanding acquired during the pre-training process[17].

b) RoBERTa:

RoBERTa (Robustly Optimized BERT Pretraining Approach, introduced by Liu et al. (2019) [5], builds on BERT's architecture by making several improvements to the pretraining process. RoBERTa has the identical transformer-based architecture as BERT but optimizes it for better performance across a variety of NLP applications, including the identification of sarcasm.

The enhancements in the architecture of RoBERTa include:

- **Training on more data:** RoBERTa uses a larger corpus that include Common Crowl dataset and for a longer time.
- **Removing the Next Sentence Prediction (NSP):** Removing does not improve performance and can be omitted without loss of contextual understanding.
- **Dynamic masking:** In RoBERTa, the tokens that are masked change dynamically during training, allowing the model to learn better contextual representations.
- **Larger batch sizes:** An increase in the length of sequences during training, enabling RoBERTa to better capture long-range dependencies between words.

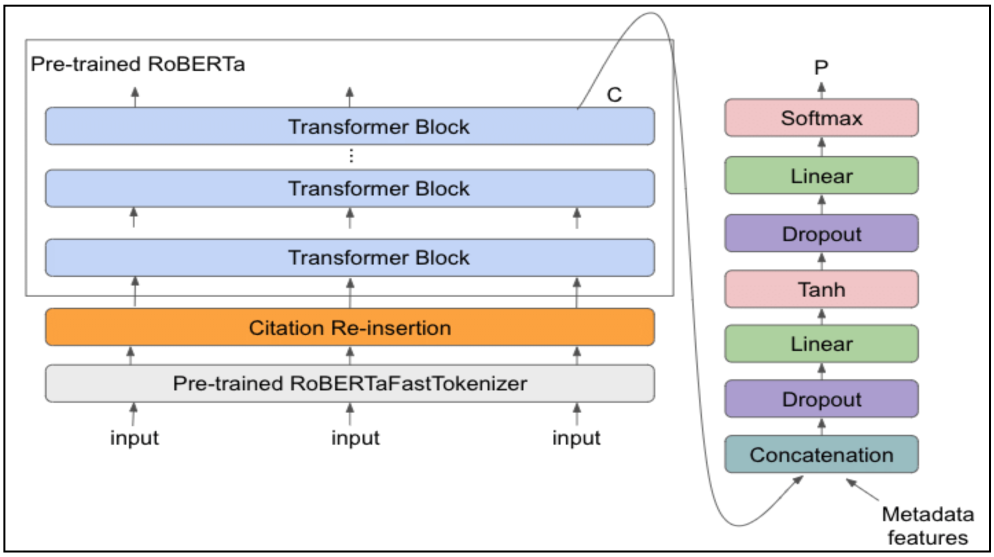


Fig. 2 Diagrammatic representation of BERT architecture [18].

In sarcasm detection, RoBERTa's ability to learn from a more extensive dataset and its dynamic masking contribute to better performance in identifying subtle cues, such as tone and irony, which are often crucial for detecting sarcasm [19].

Fine-Tuning RoBERTa

Fine-tuning RoBERTa refers to the process of taking a pre-trained RoBERTa model and adapting it to a specific downstream task by continuing its training on a smaller, task-specific dataset[29]. Fine-tuning RoBERTa involves the following steps:

- **Pre-trained Model Selection:** Start with a pre-trained RoBERTa model that has been trained on a large general corpus(RoBERTa-base in this case).
- **Task-Specific Dataset Preparation**: Obtain a labelled dataset specific to your task split it into training, validation, and test sets.
- **Model Modification**: Replace the output layer of RoBERTa with a task-specific layer suited for the number of output classes (e.g., for binary classification).

- **Fine-Tuning Process**: Train the modified model on your labelled dataset by continuing optimization of the model weights using a task-specific loss function.
- **Evaluation and Adjustment**: Validate the fine-tuned model on the validation dataset, adjust hyperparameters and monitor performance.

Dataset and Methodology :

In this section, we discuss the steps taken for preprocessing the dataset, the architecture of the models, and the specific fine-tuning process for both BERT and RoBERTa models. The models are evaluated on various metrics such as accuracy, precision, recall, F1-score, and confusion matrix to provide a detailed comparative analysis.

a) Dataset Description

For our research, we utilized the *Sarcasm Headlines Dataset* for training and testing the models which was sourced from Kaggle (Source : https://www.kaggle.com/datasets/rmisra/news-headlines-dataset-for-sarcasmdetection/data?select=Sarcasm_Headlines_Dataset_v2.json) [25].

The dataset consists of the following key attributes:

1. **"Headline":** The news headline to be classified serves as a textual input for fine-tuning the models.
2. **"is_sarcastic":** A binary label where 0 indicates a non-sarcastic headline and 1 indicates a sarcastic headline.

The dataset consists of **5,342 samples**, with the following class distribution:

- **Non-sarcastic (0):** 2,996 samples.
- **Sarcastic (1):** 2,346 samples.

b) Data Preprocessing

Before training the models, we performed the following preprocessing steps on the dataset:

1. **Tokenization**: Tokenization [27] is the process of breaking down text into smaller units called tokens, which are individual words or characters. In natural language processing (NLP), tokenization is essential for text analysis tasks such as sentiment analysis, question answering, and text classification. Both BERT and RoBERTa models require tokenization of text using their respective tokenizers [20]. We used the BertTokenizer and RobertaTokenizer from Hugging Face's Transformers library. The tokenizers convert each headline into a list of tokens and add special tokens [CLS] and [SEP] to the input sequence to mark the start and end of each headline.
2. **Padding and Truncation**: Since the input length varies, all inputs were padded to a maximum sequence length of **128 tokens** and truncated if the input length exceeded this value. This ensures uniformity across the training samples[28].
3. **Lowercasing**: All text is lowercased to maintain consistency in representation.
4. **Attention Masks**: Attention masks were created to distinguish between padded and non-padded tokens, allowing the model to focus only on the relevant tokens during fine-tuning [21].
5. **Label Encoding**: The target labels were converted into a numerical format to match the input format required for binary classification tasks.
6. **Train-Validation Split**: We split the dataset into training (80%) and validation (20%) sets to ensure the model's performance is generalizable and not limited to training data.

c) Model Fine-Tuning

For sarcasm detection, we fine-tune two transformer-based models: **BERT** (Bidirectional Encoder Representations from Transformers) and **RoBERTa** (Robustly Optimized BERT Pretraining Approach). Both models are pretrained on large corpora, and we further fine-tune them on the sarcasm detection task.

1. BERT Model

We fine-tuned the pre-trained **BERT-base-uncased** model for the task of sarcasm detection. The architecture of BERT includes 12 transformer layers, 768 hidden units, and 12 attention heads. The key aspect of BERT is its bidirectional attention, allowing it to learn from both left and right contexts simultaneously.

The final classification layer was modified to output two probabilities corresponding to the sarcastic and non-sarcastic classes [22], and the model was trained with the following hyperparameters:

- **Optimizer**: AdamW (learning rate = 2e-5).
- **Loss function**: Cross-entropy loss.
- **Batch size**: 32.
- **Number of epochs**: 3

2. RoBERTa Model

We also fine-tuned the pre-trained **RoBERTa-base** model, which is built on BERT but introduces improvements during pre-training, such as training on larger datasets and removing the next sentence prediction task. Similar to BERT, RoBERTa also includes 12 layers of transformers and 768 hidden units. The final layer was adjusted for binary classification with two output classes [23]. The hyperparameters for fine-tuning RoBERTa were set as follows:

- **Optimizer**: AdamW (learning rate = 2e-5).

- **Loss function**: Cross-entropy loss.
- **Batch size**: 32.
- **Number of epochs**: 3

Evaluation Metrics

We evaluated the models using the following metrics:

- **Accuracy**: The proportion of correctly classified samples of the model.
- **Precision**: The ratio of true positives to the sum of true and false positives.
- **Recall**: The ratio of true positives to the sum of true positives and false negatives.
- **F1-Score**: The harmonic mean of precision and recall.
- **Confusion Matrix**: A matrix displaying the distribution of predictions across the actual classes.

Result and Discussion :

a) BERT Model Results

The performance of the BERT model is summarized in the following table and confusion matrix:

Class	Precision	Recall	F1-Score	Support
Non-Sarcastic (0)	0.95	0.92	0.94	2996
Sarcastic (1)	0.91	0.94	0.92	2346

- **Accuracy**: 92.92%
- **Macro Average**: Precision = 0.93, Recall = 0.93, F1 = 0.93
- **Weighted Average**: Precision = 0.93, Recall = 0.93, F1 = 0.93

The **confusion matrix** for BERT shows that it misclassified 229 non-sarcastic headlines as sarcastic and 149 sarcastic headlines as non-sarcastic.

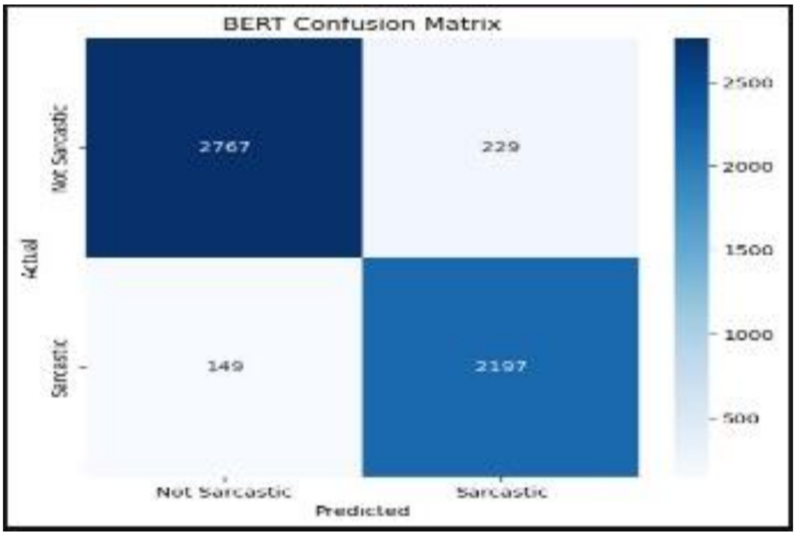


Fig. 3 Confusion Matrix for BERT Model

b) RoBERTa Model Results

RoBERTa's performance is notably poorer compared to BERT. The following table and confusion matrix provide a detailed overview:

Class	Precision	Recall	F1-Score	Support
Non-Sarcastic (0)	0.56	1.00	0.72	2996
Sarcastic (1)	0.00	0.00	0.00	2346

- **Accuracy**: 56.08%
- **Macro Average**: Precision = 0.28, Recall = 0.50, F1 = 0.36
- **Weighted Average**: Precision = 0.31, Recall = 0.56, F1 = 0.40

The **confusion matrix** reveals that RoBERTa failed to classify any of the sarcastic headlines correctly. This is a significant limitation, as the model shows a bias towards predicting non-sarcastic labels.

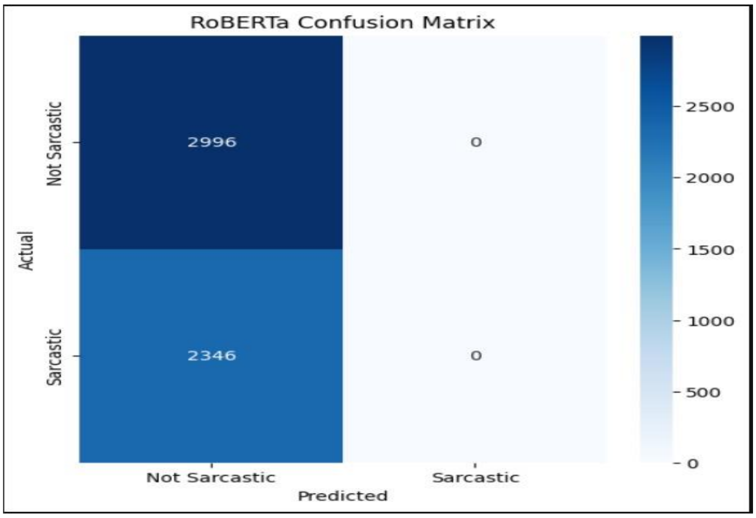


Fig. 4 Confusion Matrix for RoBERTa Model

c) Comparative Analysis

The results clearly show that BERT significantly outperforms RoBERTa in sarcasm detection. While BERT achieves an impressive accuracy of 92.92%, RoBERTa struggles with a mere 56.08%. The confusion matrix for RoBERTa shows that it classifies all headlines as non-sarcastic, which indicates that the model failed to learn useful patterns for sarcasm detection.

Several factors may contribute to this discrepancy:

1. **Fine-tuning Sensitivity**: RoBERTa may require more fine-tuning steps or adjusted hyperparameters to perform well on specific tasks such as sarcasm detection.
2. **Overfitting**: The zero precision, recall, and F1-score for the sarcastic class in RoBERTa's results suggest that the model overfitted to the majority class (non-sarcastic).
3. **Data Complexity**: The dataset contains subtle linguistic cues that BERT was better able to capture due to its bidirectional context understanding.
4. **Training Stability**: BERT's pre-trained weights and training configuration may have been better suited to this specific dataset, resulting in more stable and accurate predictions.

Conclusion :

This study conducted a detailed comparative analysis of fine-tuning BERT and RoBERTa for the task of sarcasm detection, a challenging problem in natural language processing due to the inherent ambiguity and contextual nature of sarcastic statements. The primary objective was to evaluate the effectiveness of these two pre-trained transformer models in identifying sarcastic and non-sarcastic headlines using the **Sarcasm Headlines Dataset**.

Key Findings:

1. BERT's Superior Performance:

BERT outperformed RoBERTa across all evaluation metrics. With an accuracy of **92.92%**, BERT demonstrated its ability to accurately distinguish between sarcastic and non-sarcastic headlines. The model achieved high precision, recall, and F1-scores for both sarcastic and non-sarcastic classes, showing balanced and robust performance. The bidirectional nature of BERT allows it to capture subtle contextual cues that are often necessary to understand sarcasm.

2. RoBERTa's Struggles:

RoBERTa, on the other hand, exhibited a significant drop in performance, achieving an accuracy of **56.08%**. The model failed to correctly classify any sarcastic headlines, as evidenced by its confusion matrix. RoBERTa's high recall for the non-sarcastic class and zero recall for the sarcastic class suggest that it overfitted to the majority non-sarcastic class. This may be due to various factors, including an insufficient learning rate or the inability of RoBERTa to handle sarcasm-specific patterns as effectively as BERT.

3. Contextual Understanding:

One of the most important aspects of sarcasm detection is the ability to interpret context. Sarcasm often depends on understanding both the broader context and the contradictions between different parts of a sentence. BERT's bidirectional encoder enables the model to process information in both directions (left-to-right and right-to-left), which may explain its superior ability to detect sarcasm compared to RoBERTa, which may have struggled with more nuanced context-switching.

4. Overfitting in RoBERTa:

Another critical insight is that RoBERTa's tendency to overfit to the majority class (non-sarcastic) significantly hurt its performance. This suggests that RoBERTa, in its current configuration, is not well-suited to handle imbalanced datasets where sarcastic samples are fewer in number. Adjusting hyperparameters or employing class-balancing techniques could potentially mitigate this problem, though further investigation is required.

Future Scope :

The performance gap between BERT and RoBERTa highlights several areas for future research. First, exploring different fine-tuning techniques for RoBERTa could lead to better performance, such as using different learning rates, increasing training epochs, or employing data augmentation techniques. Additionally, incorporating multimodal data (such as emojis, images, or videos) could provide valuable context that enhances sarcasm detection models. Another potential avenue for future work involves creating ensemble models that combine the strengths of both BERT and RoBERTa to achieve more balanced and generalized performance. Further studies could also explore using other variants of transformer models, such as XLNet or T5, to determine whether they can outperform BERT in sarcasm detection.

In conclusion, while BERT has proven to be a more effective model for sarcasm detection in this study, there remains significant potential for improving RoBERTa's performance and further enhancing sarcasm detection techniques in NLP.

REFERENCES :

[1] Shu, Xin. (2024). BERT and RoBERTa for Sarcasm Detection: Optimizing Performance through Advanced Fine-tuning. Applied and Computational Engineering. 97. 1-11. 10.54254/2755-2721/97/20241354.

[2] H. Yaghoobian, H. R. Arabnia, and K. Rasheed, "Sarcasm Detection: A Comparative Study," 2021.

[3] Hazarika, D., Poria, S., Gorantla, S., Cambria, E., Zimmermann, R., Mihalcea, R. (2018). Cascade: Contextual sarcasm detection in online discussion forums. arXiv preprint arXiv:1805.06413..

[4] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

[5] Liu, Y., Ott, M., Goyal, N., et al. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach.

[6] Dadu, T., Pant, K. (2020). Sarcasm detection using context separators in online discourse. In Proceedings of the Second Workshop on Figurative Language Processing (pp. 51-55)

[7] Mao, J., Liu, W. (2019). A BERT-based Approach for Automatic Humor Detection and Scoring. IberLEF@ SEPLN. 2421:197-202.

[8] Mangal, D., & Makwana, H. (2025). Performance analysis of different BERT implementation for event burst detection from social media text. *Indonesian Journal of Electrical Engineering and Computer Science*, *38*(1), 439. https://doi.org/10.11591/ijeecs.v38.i1.pp439-446

[9] Karkiner, Z., & Sert, M. (2024). *Sarcasm Detection in News Headlines with Deep Learning*. https://doi.org/10.1109/siu61531.2024.10600986

[10] Ou, L., & Li, Z. (2025). Multi-modal Sarcasm Detection on Social Media via Multi-Granularity Information Fusion. *ACM Transactions on Multimedia Computing, Communications, and Applications*. https://doi.org/10.1145/3715139

[11] Shi, C. (2025). Research Advanced in Sarcastic Detection based on Deep Learning. *Theoretical and Natural Science*, *79*(1), 23–27. https://doi.org/10.54254/2753-8818/2025.19930

[12] Hossain, Md. M., Hossain, M. J., Mridha, M. F., Safran, M., & Alfarhood, S. (2025). Multi task opinion enhanced hybrid BERT model for mental health analysis. *Dental Science Reports*, *15*(1). https://doi.org/10.1038/s41598-025-86124-6

[13] Wang, S., Liu, Q., Hu, Y., & Liu, H. (2025). Public Opinion Evolution Based on the Two-Dimensional Theory of Emotion and Top2Vec-RoBERTa. *Symmetry*, *17*(2), 190. https://doi.org/10.3390/sym17020190

[14] Garcia, E., Felipe da Silva, N., Gomes, J., Albuquerque, H., Souza, E., Siqueira, F., Lima, E., & Pacheco de Carvalho, A. A. (2024). RoBERTaLexPT: um modelo RoBERTa jurídico pré-treinado com deduplicação para língua Portuguesa. *Linguamática*, *16*(2), 183–200. https://doi.org/10.21814/lm.16.2.457

[15] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*. https://arxiv.org/abs/1706.03762

[16] Dutta, S., Chakraborty, A. (2019). A deep learning-inspired method for social media satire detection. In: *Soft Computing and Signal Processing* (pp. 243–251). Springer.

[17] Wang, Z. (2023). *A new computationally efficient method to tune BERT networks – transfer learning*. *2580*. https://doi.org/10.1088/1742-6596/2580/1/012012

[18] https://images.app.goo.gl/hGd5DYZZzNf7BLnZ6 (Roberta image)

[19] Wu, P., Zong, J., & Zhou, Z. (2021). Exploring RoBERTa for Sarcasm Detection on Social Media. *Journal of Computer Science and Technology*, 36(1), 123-135.

[20]Kumar, A., Sharma, A., & Agarwal, A. (2020). BERT-Based Sarcasm Detection on Social Media: A Multimodal Approach. *Proceedings of the 28th International Conference on Computational Linguistics (COLING)*.

[21] *Finetuning for Sarcasm Detection with a Pruned Dataset*. (2022). https://doi.org/10.48550/arxiv.2212.12213

[22] Handoyo, A. T., Diot, A., Suhartono, D., & Lamiroy, B. (2023). Measuring the Quality of Semantic Data Augmentation for Sarcasm Detection. *International Journal of Intelligent Engineering and Systems*, *16*(5), 79–91. https://doi.org/10.22266/ijies2023.1031.08

[23]Youwei Song, Jiahai Wang, Tao Jiang, Zhiyue Liu, and Yanghui Rao. 2019. Attentional encoder network for targeted sentiment classification. arXiv preprint arXiv:1902.09314.

[25]https://www.kaggle.com/datasets/rmisra/news-headlines-datasetforsarcasmdetection/data?select=Sarcasm_Headlines_Dataset_v2.json

[26] https://images.app.goo.gl/yK4HFJuuBjn135Sk6 (Bert image)

[27] Rolandos Alexandros Potamias, Georgios Siolas, and Andreas-Georgios Stafylopatis. 2020. A transformer-based approach to irony and sarcasm detection. Neural Computing and Applications, 32(23):17309–17320.

[28] Webster, Jonathan J., and Chunyu Kit. "Tokenization as the initial phase in NLP." COLING 1992 volume 4: The 14th international conference on computational linguistics. 1992.

[29] He, L. (2024). Enhanced twitter sentiment analysis with dual joint classifier integrating RoBERTa and BERT architectures. *Frontiers in Physics*, *12*. https://doi.org/10.3389/fphy.2024.1477714