Improving Sentiment Analysis Accuracy: A Novel Deep Learning Approach to Sarcasm Detection

1. Abstract

Sentiment Analysis (SA) plays a critical role in understanding public opinion, decision-making processes, and consumer preferences. However, the detection of sarcasm within sentiment-laden text poses a significant challenge due to its inherent complexity and ambiguity. Sarcasm often involves context, implicit meanings, or tone that traditional models fail to capture effectively. This research proposes a novel deep learning-based framework for sentiment analysis integrated with sarcasm detection to improve the accuracy and robustness of sentiment prediction.
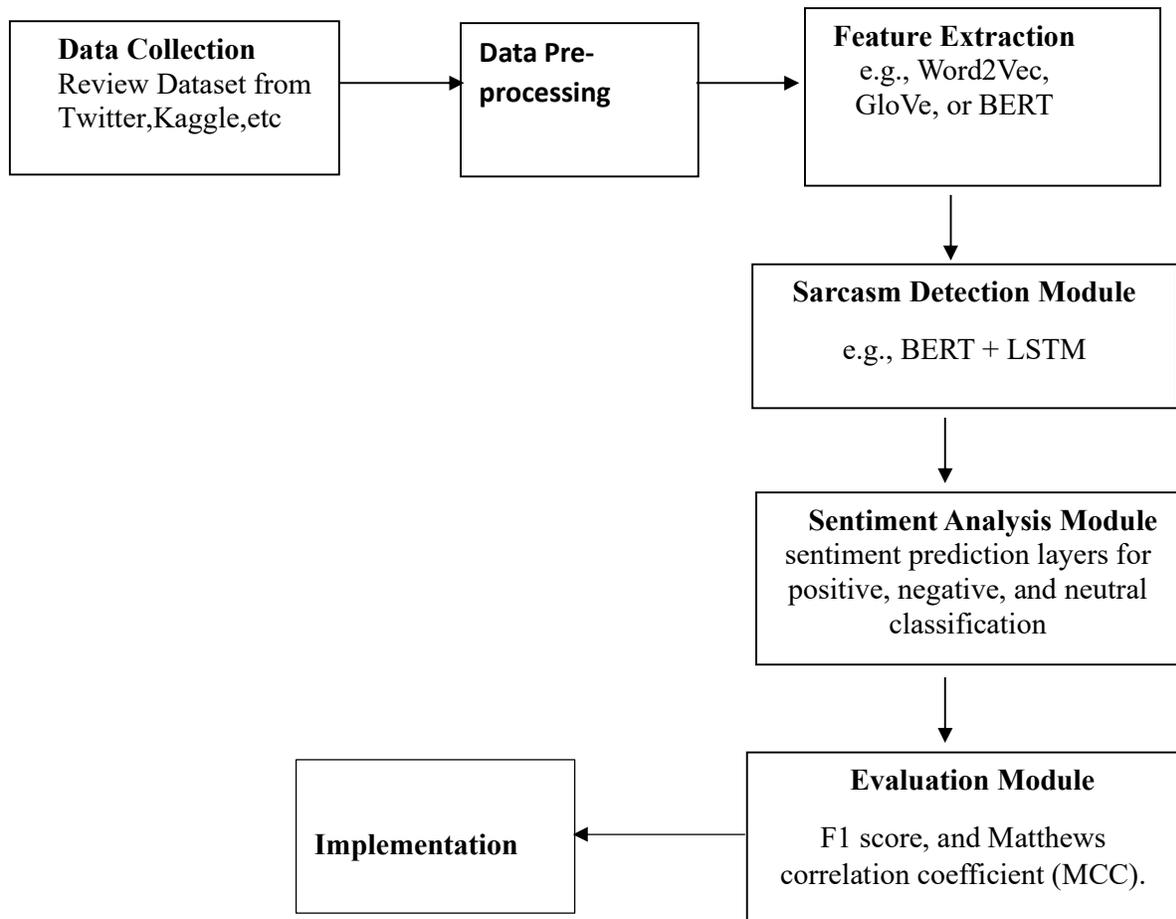
The model will leverage advancements in Natural Language Processing (NLP) and deep learning techniques, focusing on context-aware mechanisms, linguistic feature extraction, and multimodal analysis for enhanced sarcasm detection. The proposed framework aims to bridge the gap between conventional sentiment analysis models and the nuanced understanding of sarcastic expressions.

2. Objectives

1. **Develop a robust deep learning model** for sentiment analysis with integrated sarcasm detection capabilities.

2. **Explore context-aware architectures** (e.g., transformers, attention mechanisms) to improve sarcasm detection accuracy.

3. **Incorporate linguistic and semantic features** such as humor, irony, and ambiguity to enhance sentiment interpretation.

4. **Analyze multimodal data** (text, audio, and image inputs) to evaluate the interplay of sarcasm in diverse formats.

3. Methodology

1. **Data Sources**: Twitter, Reddit, Amazon reviews, Kaggle.

2. **Tools and Frameworks**: TensorFlow, PyTorch, HuggingFace Transformers, NLTK, and OpenCV for preprocessing and model development.

3. **Model Development:** Leverage advanced deep learning architectures such as transformers (e.g., BERT, GPT) to capture semantic and contextual information.

4. **Model Training**: Train on labeled datasets with a focus on sarcasm-labeled data. Utilize transfer learning for efficient training.

5. **Model Validation**: Validate using k-fold cross-validation and benchmark against state-of-the-art models such as BERT, RoBERTa, and GPT.

6. **Implementation in Sentiment Analysis:** Integrate the sarcasm detection model into a sentiment analysis system and measure its impact on overall accuracy.

```
┌─────────────────────┐      ┌──────────────┐      ┌─────────────────────┐
│   Data Collection   │─────▶│ Data Pre-    │─────▶│  Feature Extraction │
│ Review Dataset from │      │ processing   │      │    e.g., Word2Vec,  │
│ Twitter,Kaggle,etc  │      │              │      │    GloVe, or BERT    │
└─────────────────────┘      └──────────────┘      └─────────────────────┘
                                                              │
                                                              ▼
                                          ┌─────────────────────────────┐
                                          │  Sarcasm Detection Module    │
                                          │                              │
                                          │     e.g., BERT + LSTM        │
                                          └─────────────────────────────┘
                                                              │
                                                              ▼
                                          ┌─────────────────────────────┐
                                          │  Sentiment Analysis Module   │
                                          │ sentiment prediction layers  │
                                          │ for positive, negative, and  │
                                          │    neutral classification    │
                                          └─────────────────────────────┘
                                                              │
                                                              ▼
    ┌─────────────────────┐              ┌─────────────────────────────┐
    │                     │◀─────────────│     Evaluation Module        │
    │   Implementation    │              │  F1 score, and Matthews      │
    │                     │              │ correlation coefficient (MCC)│
    └─────────────────────┘              └─────────────────────────────┘
```

4. Expected Outcomes

1. A high-accuracy sentiment analysis model capable of detecting sarcasm with context-aware mechanisms.

2. Improved sentiment prediction accuracy for ambiguous or sarcasm-laden text.

3. A publicly available dataset and model for sentiment analysis with sarcasm detection for researchers.

5.References

1. A. D. Yacoub, S. Slim, and A. . Aboutabl, "A Survey of Sentiment Analysis and Sarcasm Detection: Challenges, Techniques, and Trends", IJECES, vol. 15, no. 1, pp. 69-78, Jan. 2024

2. Razali, M. S., Halin, A. A., Ye, L., Doraisamy, S., & Norowi, N. M. (2021). Sarcasm detection using deep learning with contextual features. *IEEE Access*, *9*, 68609-68618

3. Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, *470*, 443-456