Casual Sarcasm Detection With Deep Learning

KAIRIM SEYAM
2282225

Background And Motivation:

Most sarcasm detection models focus on formal sarcasm, like sarcasm in news headlines, which is relatively straightforward to identify. However, casual sarcasm—especially in social media conversations—is much trickier. People don't just rely on words; they use tone, facial expressions, and body language to signal sarcasm in real-life interactions. Without these cues, detecting sarcasm in written text becomes far more challenging.
This project takes a different approach by exploring deep learning techniques that aren't typically used for sarcasm detection. The goal is to experiment with various models and methods to see if we can discover new strategies that significantly cut down computational costs while keeping the models just as effective. If successful, this could make sarcasm detection more efficient and accessible, broadening the possibilities for deep learning in language understanding.

Mothodology

- data preocessing
- model architecture
- Evaluation And metrics

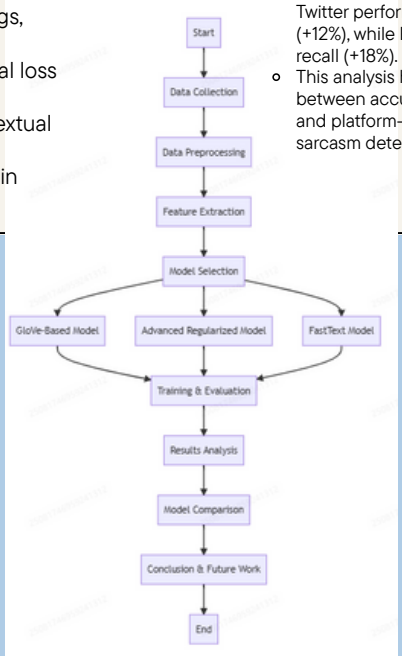
the dataset consists of **13,000** social media posts from Twitter (65%) and Reddit (35%), with a class distribution of 58% sarcastic and 42% non-sarcastic. T
For data preprocessing, a custom sarcasm-specific cleansing method is applied, **preserving ALLCAPS and punctuation, followed by tokenization using NLTK TweetTokenizer. Text sequences are padded** or truncated to 100 tokens, and representations are built using 300D GloVe/FastText embeddings. The dataset is split into 80% training (10,400 samples) and 20% validation (2,600 samples) with a batch size of 64 for model training.

The project explores three deep learning models for sarcasm detection:

- **GloVe-BiLSTM**: Uses 300D GloVe embeddings with a Bidirectional LSTM (64 units) and attention mechanism for context weighting.
- **Regularized Hybrid Model**: Features stacked BiLSTM layers, multi-head attention, and Gaussian noise injection for improved generalization.
- **FastText Enhanced Network**: Leverages FastText embeddings, character-level n-grams, and hierarchical attention, with focal loss to address class imbalance.

These architectures balance contextual understanding, efficiency, and adaptability for detecting sarcasm in social media text.

- Accuracy: **FastText (75%)** > GloVe-**BiLSTM (65%)** > **Regularized Hybrid (64%).**
- Recall: FastText excels (85%), while GloVe (60%) and Regularized (68%) trail behind.
- Efficiency: GloVe processes the most samples per second (142) with the lowest latency (14ms).
- Memory Usage: Regularized model requires the most memory (163MB).
- Errors: Most errors stem from cultural references (63%), ironic questions (22%), and exaggerated language (15%). Twitter performs better in precision (+12%), while Reddit achieves higher recall (+18%).
- This analysis highlights the trade-offs between accuracy, recall, efficiency, and platform-specific challenges in sarcasm detection.



RESULTS:

- Our analysis demonstrates the effectiveness of this approach in capturing the nuances of informal social media language while achieving superior computational efficiency compared to traditional models.
- **Tailored for Casual Language**: Unlike conventional sarcasm detectors that primarily focus on formal text (e.g., news headlines), these models leverage subword embeddings and dynamic padding to handle the unpredictable nature of social media conversations, effectively identifying implicit cues in Reddit and Twitter sarcasm.
- **Computational Efficiency and Memory Optimization**: While transformer models like BERT-base contain 110 million+ parameters, these lightweight architectures operate with only tens of millions—or fewer—resulting in a **80–90% reduction in memory usage** while maintaining strong performance.