Sarcasm Detection in Social Media Comments Using NLP

Abstract:
Sarcasm in online comments presents a challenge for sentiment analysis systems, as it can distort the true meaning behind a statement. This project aims to build a **sarcasm detection system** using **Natural Language Processing (NLP)** to analyze social media comments.

Methodology:
The system will collect social media comments and preprocess the text using **text normalization**, **tokenization**, and **word embeddings** (e.g., **Word2Vec** or **GloVe**). To detect sarcasm, the system will leverage **contextual analysis** through models like **BERT**, which understand not only individual word meanings but also the broader context of a sentence. A combination of **sentiment analysis**, **emotion detection**, and **textual cues** (e.g., exaggerated language, contradictory sentiments) will be used to identify sarcastic remarks. The model will be trained using labeled datasets containing sarcastic and non-sarcastic examples.

Outcome:
The expected outcome is a high-accuracy sarcasm detection system that can discern sarcasm in online comments and social media posts. This will improve sentiment analysis systems by providing more accurate representations of user opinions, especially in cases where sarcasm might mislead traditional sentiment classifiers. Additionally, it will enhance the quality of automated content moderation and sentiment reporting tools.