**Sentiment Analysis of IMDb Data: A Problem Statement and Literature Review**

**Problem Statement:**

The rapid expansion of digital media has precipitated a surge in user-generated content, particularly evident on platforms like IMDb where reviews about movies and TV shows generate. Examining the sentiments expressed within these reviews is very helpful in considerate audience perspectives. Sentiment analysis, a facet of natural language processing (NLP), endeavours to automatically identify and extract subjective information from text to recognise prevailing sentiments.

However, conducting sentiment analysis on IMDb data poses multifaceted challenges. Primarily, IMDb harbours a diverse array of data, enclose reviews of disparate lengths, writing styles, and languages. Moreover, the prerequisite for labelled data to train machine learning models, integral for sentiment analysis, may be deficient or unreliable within the context of IMDb reviews. Additionally, the presence of sarcasm, irony, and nuanced expressions sharpen the complexity of sentiment classification. Hence, crafting precise and resilient sentiment analysis models tailored to IMDb data is a formidable yet essential endeavour.