Team Members:
1. Arun Gaonkar (unity ID: agaonka)
2. Manasi Sanjay Ghosalkar (unity ID: mghosal)

Context Based Sarcasm Detection

Sarcasm detection is an important aspect of natural language understanding in a social context. A sentence can essentially be interpreted in an entirely opposite manner than intended if not looked at in reference to the prior conversation or existing knowledge. This makes it highly contextual. Sarcasm detection is closely related to sentiment analysis in the sense that the sentiment of text depends on the context and situation, similar to sarcasm. So it is important to detect sarcasm in texts to understand the exact meaning the person is conveying.

We are proposing an approach for detecting sarcasm in sentences based on prior context. For this, we plan to use a News Headlines dataset which has labelled sarcastic and non-sarcastic headlines along with the URLs to the corresponding news article. The content of the article will first be extracted to be used as context for the headlines. It will then be used to predict the sarcasm in the headline.

We are getting the dataset from Kaggle which has around 28000 rows.
Dataset Link:
https://www.kaggle.com/datasets/rmisra/news-headlines-dataset-for-sarcasm-detection

Initial planning is to use different embeddings (like word2vec, paragraph2vec) to extract the contextual meaning and then build models for classification. We can also include entities (after extracting it using n-grams or NER) of each sentence and compare the prediction results.

For the model, one approach that we want to try is to treat sarcasm detection as an extension to sentiment analysis. Layers from a model pre-trained for sentiment analysis can be appended with additional layers and then fine-tuned for the sarcasm detection task.