**RoBERTa for Sentiment Analysis**

**Harsha Vardhan Khurdula**

**7th of September, 2023.**

**Introduction**

The current task at hand is to use Transformers in-order to determine which of them can be used for our Sarcasm dataset aggregated from Reddit.

**Approach**

In my previous report, I had mentioned “Data is Fundamental” for this project/research. And I stick by it. The dataset which we worked on previously was Twitter based dataset. I was of the opinion to try and explore where things go with a different dataset. Since my partner Mr. Naik had already performed some experimentation with the Twitter based dataset, I wanted to see how RoBERTa would work against Reddit based sentiment dataset all trained in English.

Keeping this in mind, I tested this dataset against some versions of RoBERTa:

1. SieBERT – English Language – Large Model:

The name SieBERT stands for Sentiment in English BERT. It is a fine-tuned checkpoint of RoBERTa-large. It performs binary classification given a statement to determine the sentiment behind it.