**Classifying Tweets Based on Climate Change**

By Jonnada Sai Rahul

**Problem Statement**

Climate change has become an increasingly polarizing topic in popular and political media. So in this project we develop a classifier that can discriminate between text that shows belief vs disbelief in human caused climate change.

So here comes the process:

1. **Import Libraries and packages:**

Importing all the required Libraries and packages for the given problem.

1. **Importing the Dataset:**

We are required to import the dataset which is very important to proceed further. Dataset can be in any format like CSV, JSON, EXCEL etc.,

1. **Data Pre-processing:**

Making a raw data into useful and understandable like we can keep only the data which we require and can drop which is not useful which makes data clean and easy access

This includes Data Cleaning, Data Transformation, Data Reduction.

In Problem is related only to the sentiment [-1,0,1] but our data consists of extra data i.e. [2] so we can drop that. This is an example of Pre-processing.

We should remove all unwanted symbols which may cause a problem while doing the task.

1. **Method we use (Algo):**

Multinomial Naive Bayes (MNB):

1. MNB Unigram
2. MNB Bigram
3. **Training the Model:**

Train the model with the data we have imported after the pre-processing.

1. **Testing Model:**

Finally, we test our model working (Accuracy) check the error percentage and test how the model works for other data which is not in our dataset.