**Fake -News Detection Project**

**Problem statement:**

The authenticity of Information has become a longstanding issue affecting businesses and society, both for printed and digital media. On social networks, the reach and effects of information spread occur at such a fast pace and so amplified that distorted, inaccurate, or false information acquires a tremendous potential to cause real-world impacts, within minutes, for millions of users. Recently, several public concerns about this problem and some approaches to mitigate the problem were expressed.

In this project, you are given a dataset in the fake-news\_data.zip folder. The folder contains a CSV files train\_news.csv and you have to use the train\_news.csv data to build a model to predict whether a news is fake or not fake. You have to try out different models on the dataset, evaluate their performance, and finally report the best model you got on the data and its performance.

**Data- Description:**

There are 6 columns in the dataset provided to you. The description of each of the column is given below:

“id”: Unique id of each news article

“headline”: It is the title of the news.

“news”: It contains the full text of the news article

“Unnamed:0”: It is a serial number

“written\_by”: It represents the author of the news article

“label”: It tells whether the news is fake (1) or not fake (0).

**Submission:**

You have to submit the jupyter notebook, in which you have built your best performing model. Your jupyter notebook should be well commented so that it can be easily understood what are you actually trying to do in the code. Also mention which is your best performing model and the measure of its performance (accuracy score , f1 score etc)