**Fake News Detection Project Proposal**

* **Introduction:** Fake news spread misinformation, disinformation or mal-information. They spread through word of mouth, traditional media, digital forms of communication such as edited videos, memes, unverified advertisements and social media propagated rumors. They have potential of resulting in mob violence, suicides, people following misinformation during pandemics, panic spreading, stampedes etc. I came across a dataset at Kaggle which can be used to develop models based on natural language processing to detect fake news. Developing models to develop approaches to detect fake news can help in detecting what is fake and true. It can save a lot of financial hazards, social and miscellaneous havocs. This dataset contains 40000 articles consisting of fake as well as real news. The fake and real news data is given in two separate datasets with each dataset consisting around 20000 articles each.
* **Hypothesis:** Create and train a machine learning model so that it can correctly predict whether a given piece of news is Real or Fake
* **Criteria for success**: This project proposes the question of whether it is possible to detect fake news through machine learning models. Specifically, the aim of this project is to determine the ideal model that is efficient in predicting fake news while also limiting the cost of memory and storage for computation
* **Challenges:** Feature extraction, right preprocessing of the text, choosing the right models and comparing the accuracy