Fake News Detection

**Group Name**: Dhumgod

CB.EN.U4ELC19004 – Aditya Sampath

CB.EN.U4ELC19029 – Manit Maini

CB.EN.U4ELC19036 – Pradhumna Guruprasad

CB.EN.U4ELC19052 – Shreyas Nagesh

# Abstract

News tweets that report what is happening have become an important real-time information source. Incidents happening in different corners of the globe are first reported by individuals on the ground than the media houses. This is where the issue arises. We cannot be sure how many of these tweets are true and how many are fake. This project proposes to use Semantic Role Labelling (SRL) to help classify and categorize tweets based on whether they are true or not. A domain specific SRL system will be trained with a supervised learning approach to resolve the problem.

# Introduction

Fake news is becoming more prevalent in today's political scene. Researchers and media professionals have developed fake news detectors that use natural language processing (NLP) to examine word patterns and statistical correlations in news items to assist solve this issue. While these detectors achieve amazing accuracy on known examples of edited news, the analysis is often fairly shallow. Models examine whether news pieces correspond to normal standards and styles employed by experienced journalists. This has two consequences. First, these algorithms can detect false news only when they are poorly written, such as when the substance is completely unrelated to the title or when the article contains terms that are regarded prejudiced or provocative. While this criterion is sufficient to detect many existing cases of fake news, more competent rumour disseminators can construct more subtle attacks, such as interfering with a well-written actual news piece in a targeted manner. An antagonistic article can readily elude discovery by retaining the original subject matter and connecting the content firmly to the title without employing biased words. Through our project we strive to solve the shortcomings of the previous described model for Hindi language.

# Methodology

We first collected a labelled dataset of tweets in English denoting whether each tweet is fake or true. As we can’t directly use this data in our ML model, we first preprocess the text so that it makes more sense to the algorithm. First tokenization was done, followed by stop word removal. This was followed by lemmatizing the data so that we extract only the meaningful part of the word. This is followed by POS tagging which adds additional details to each word.

Once this preprocessing stage is over, we propose to use a bidirectional language model like BERT to bring out the best correlations between word embeddings and give a greater accuracy than a simple bag of words or n-gram language model.