Real OR Fake News Analysis

This dataset consists of about 40000 articles consisting of fake as well as real news. Our aim is train our model so that it can correctly predict whether a given piece of news is real or fake. The fake and real news data is given in two separate datasets with each dataset consisting around 20000 articles each.

Business Objective:

1. Need to classify the fake and real news accurately.

Architecture level analysis:

1. Data transformation/Text processing using R/Python
2. Need to get sentiments Analysis and n-gram analysis with some charts like histogram, Density plot, Barplot, pie-plot etc.
3. Deployment through R Shiny or Flask/ Streamlit

Milestones: Phase 1

Milestone	Duration	Task start - End Date
Kick off and Business Objective discussion	1 day	23-Nov-2024
Data set Details	1 Week – 1 ½ week	depends
EDA	1 Weeks – 1 ½ week	30-Nov-2024
Model Building	1 Week – 1 ½ week	7-Dec-2024
Model Evaluation	1 week	14-Dec-2024
Feedback		
Deployment	1 Week	21-Dec-2024
Final presentation	1 day	depends

Protocols:

1) All participants should add here to agreed timelines and timelines will not be extended
2) All the documentation – Final presentation and R/python code to be submitted before the final presentation day
3) All the participants must attend review meetings