

Capstone Project 2 : Project Proposal

Project Title :

“Fake News Classifier”

Project Mentor : Vikas

What is the Problem to be solved?

In Daily routine we will read a number of articles about trends , politics , technology, food etc., but we just read and will come to our opinion and we will not think whether that article is Fake or Real

Who is this project beneficial for?

This project is used for all of us who have a habit of reading news articles. So this project will help us to identify a given article is Fake or Real

So What is the data used for the project? And where I received the data from?

So the data consists of all News records about different topics in the world and I have extracted the data from Kaggle website. This is a multivariate data consisting of Id, title, author, text, label, etc.

Techniques that I would use to solve the problem:

- 1) Data Wrangling - I will observe the data for all the missing values, incorrect entries and based on the observation I will clean the data for further use and apply the Modern NLP techniques to transform text data into vector form
- 2) Exploratory Data Analysis - I would plot the relation between different parameters in the dataset and based on that I can find the trends for storytelling.
- 3) Machine Learning Model / Deep Learning - After all the initial work on the dataset it would be time to build a Machine Learning Model to find the given article is Fake or Real i may use the LSTM or a simple Naive Bayes.

What is the methods I will use to deliver the project?

I will be using the following tools to work on the project :-

- Jupyter Notebook for the major part of the project. - I will be using Python version 3.x for writing the code. - I will use Google doc for submission on the in process project reports. - Powerpoint for Data Storytelling presentation.

References:

- Springboard Curriculum material - DataCamp courses - Google, Stackoverflow, Github, etc. - And any other Open Resources.