**CKME -136: Data Analytics - Capstone Project Abstract**

***Research problem and Theme***

The rising issue of spreading fake information in all social media has a great impact on individuals and society. Exclusively in politics, social media is played an important role that influences voters in 2016 presidential election. Detecting real or fake news becomes very necessary and challenging and has tremendous attention. The theme chosen for this project is classification algorithm.

***Research question and solution***

How will you make an algorithm to determine if an article is fake or not?

The solution is to build a classification algorithm that can detect and identify accurately if a piece of article is real or fake news. By achieving, it would be beneficial to help in determining which is real or fake news.

***Source of the data***

The dataset obtained from <https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset>

which includes True and Fake files, each has four columns, rows of True is 21417 and Fake is 23538. The data based on political and world news and collected for analysing from January 2016 to December 2017.

***Technique***

My initial experiment will be carried out with NLP, and widely used machine learning methods such as Text Classification, Support Vector Machine and Multinomial Naïve Bayes and planning to go in detail for making better predictions and accuracy improvement by using LSTM deep learning method. Evaluating classification performance measure of my model via accuracy, F1 score and AUC-ROC.

***Tool***

I will be using “R and Python” language for showcasing my findings.