# WEEK 4

## Resources

**This week will have 2 minor problem statements.Please refer to these links below for tutorials needed to approach this week's PS:**

- A blog on basic ideas behind sentiment analysis: https://towardsdatascience.com/a-step-by-step-tutorial-for-conducting-sentiment-analysis-a7190a444366
- LSTM explained: https://youtu.be/QciIcRxJvsM
- LSTM tutorial (Pytorch): https://youtu.be/AvKSPZ7oyVg
- LSTM + Rolling Window algorithm for future prediction (easy to use reference notebook using Tf.keras): https://github.com/shubhambhalala/Live_Stock_Market_Forecasting/blob/master/Live_Stock_Prediction.ipynb

---

### Assignment 1 : Oil Price Prediction

**In this problem, a single column of Oil Prices will be provided to you.The task is to predict the oil price for the next 30 days using Univariate LSTM neural network.The evaluation metric for this PS would be RMSE , given price prediction is essentially a regression problem**.

Link to the dataset (Only use the Date and the Price columns and ignore the rest of the columns):
https://drive.google.com/file/d/1lKLs-uyQ85JkoXhk-Hwom5q9RRn9Fbd4/view?usp=share_link

Task : Predict future prices for the next 30 days
Evaluation Metric: RMSE

---

### Assignment 2:Sentiment Analysis

In this problem, you will be provided with a text dataset and the aim is to use the NLTK Library to determine the **Sentiment** of the text in the **Sentence** column.

- Link to the dataset: https://drive.google.com/drive/folders/17VUgOJ_2dO6vSdVgisXQEc1_OtUf-9XU?usp=share_link
- Make sure to follow the stepwise order of operations (tokenization,stopwords, lemmatization etc)
- Evaluation metric: f1_score.