IND5003

**Project Idea:**

Sentiment Analysis

**Main Aim:**

Predict market trend based on what was posted on Stocktwit about Apple

**Process:**

1. Extract Data from Stocktwit using APIs - (date, body)
2. Pre-processing of Data
3. Sentiment Analysis
   1. Generate a sentiment score for the posts that comes out each day
4. Plot over time and compare against stock trends
5. (optional): Use sentiment score time series to predict market trend with machine learning

**Sentiment Analysis Model**

Naïve Bayes

Unsupervised classification into positive or negative classes

Or use pre-trained model – TextBlob: Naïve Bayes and pattern analyser

Deep learning

Gives a score from 0 to 1 (need to train with labelled data – could manually score data upon 5)

Could find pre-trained models

VADER

Pre-trained

Gives a score from -1 to 1, manually check and edit the labels and use data + edited sentiment scores as training data for deep learning

Deep Recursive Trees

Pre-trained RNN

Google API

Aylien API – document level sentiment analysis

Indico API

Compare models based on accuracy (manually label tweets and compare with results given by models)

Creating feature:

Average all sentiments in the period (1 day, or split the day into parts)

Variance of the distribution of sentiments over the period

Give weight to sentiment based on number of likes

**Predicting stock price**

Deep learning

Feature include sentiment score time series and stock price time series in windows

Could look at percentage change from the open to close ((close-open) / open)