**Springboard Data Science Career Track 2020**

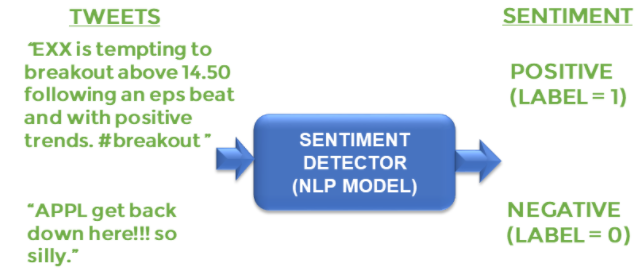
**Capstone Project 2: Project Proposal**

**Stocks Sentiment Analysis Using AI**

**Mukkul Kurwa**

**September 2020**

**A problem statement —** We live in a world where we are constantly bombarded with social media feeds, tweets, and news articles. This huge data could be leveraged to predict people sentiment towards a particular company or stock. Natural language processing (NLP) works by converting words (text) into numbers. These numbers are then used to train an Al/ML model to make predictions. Al/ML-based sentiment analysis models, can be used to understand the sentiment from public tweets, which could be used as a factor while making a buy/sell decision of securities.



**Data—** The data is already collected from Twitter and has been labelled. I’ll clean the data by removing punctuations and stop words. I’ll then visualize the data. Finally, the data will be prepared by applying tokenization and padding.

A **solution statement** — I’ll be using RNN and LSTM in order to build a custom-based deep neural network to perform sentiment analysis. It will just help us decide whether to buy or sell securities.

A set of **evaluation metrics** — Once the model is trained, I need to test its performance on the testing dataset. The model has never seen this information before; as a result, the testing dataset allows me to determine whether or not the model will be able to generalize to information that wasn't used during its training phase. I will use some of the metrics for this purpose such as confusion matrix.