# Project: Impact of Twitter Feed on Stock Price Predictions

This project explores and analyzes the impact of Twitter Feed on Stock Price Prediction. The Stock price prediction is performed using Long Short-Term Memory Neural Network. This is a flavor of Recurrent Neural Networks specifically designed to train models with sequential data like stock prices. We will use the Keras version of LSTM Neural Network.

The data for training the model will be taken using an API provided by Alpha Vantage (<https://www.alphavantage.co/>). This API provides real-time and historical data for any stock. We will use historical data for training and testing. The open values for each day will be taken to predict the stock prices. The API also provides a way to save the data into a CSV file which makes it convenient to load it into a Dataframe.

This project contains three Jupyter notebooks which server the following purpose:

1. **StockPricePredictionWithLSTM.ipynb:** This will fetch the stock data using the API. The data will then be normalized and split into test and train data. Then will be prepped and bundled for inputting to the LSTM for fitting the model. We will use the test data to check the predictions vs actual values. This model will serve as the original model and will be compared with the next model which will contain the twitter data as input.
2. **TwitterSentimentalAnalysis.ipynb:** This notebook will scrape for twitter feeds. We will not be using the API provided by Twitter as it has limitations to fetch data only till past 7 days. So, we will be using GetOldTweets3 library to scrape. We will use the time stamp from the stock data as a search parameter for the tweet scraper. For each day we will take a sample of tweets. For each of the tweet in the sample, we will get the sentimental polarity using TextBlob library and aggregate the score which represents the score for that timestamp. We will save all the data scraped and the score in a CSV. And this will be used in the next Jupyter notebook
3. **TwitterAndStockDataTogether.ipynb:** This notebook integrates the data scraped from the previous notebook and its score with a new LSTM model. The LSTM model will be built and trained in the same way as the original model but with the twitter score. Once the model is trained, we get to calculate the residuals and compare with the original model. We analyze the final model and the twitter data to understand the impact of twitter feed on the stock price prediction.

**Conclusion**:

The overall analysis of the tweets and their impact show us that even though sentimental analysis score of twitter feeds does help in predicting the stock prices, the twitter feeds might not be a very reliable source of information which can be used to predict future prices. This is because mostly twitter feeds contain data about what has already happened. Selecting the relevant tweets from the feed might be a solution that could help identify a trend. But that would require us to build a model to scrape through all the tweets and selecting only those tweets that are relevant. For this project and notebook, we took samples from twitter feed and used them for analysis.

For future scope, we can develop a Tweets Classifier (maybe a Neural network) that could possibly select tweets that are relevant to predicting stocks (This classifier would be similar to a spam filter which selects the important messages).

**Contact:**

LinkedIn**:** <https://www.linkedin.com/in/ninadsubhedar/>

GitHub: <https://github.com/ninadsubhedar>

**License:**

MIT License

Copyright (c) 2019 Ninad Subhedar

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.