**Description**

News media has a tremendous effect on the stock market prices Financial news articles are read every day by millions, interpreted by companies, and analyzed by automated stock algorithms at quantitative trading firms. Key news and data about company trades, revenues, market climates, and geopolitical conflicts strongly influence trading and market transactions. In our project, we address the problem of predicting stock price trends using news headlines. We want to determine the extent to which we can predict the movement of stock prices using features extracted from online news.

Related Work Stock market prediction is an area of very active research. Attempts have been made to predict stock prices using technical information such as company technical and price history Recent work has begun to incorporate natural language text data into prediction – for example, using full-text of news articles hidden Markov model incorporating news headlines and sentiment in social media . Most of these works has worked with short term price movements (on the order of minutes), while we will focus on end of day stock prices; we find that a longer time horizon is less susceptible to noise, and more beneficial to the common investor

**Mission**

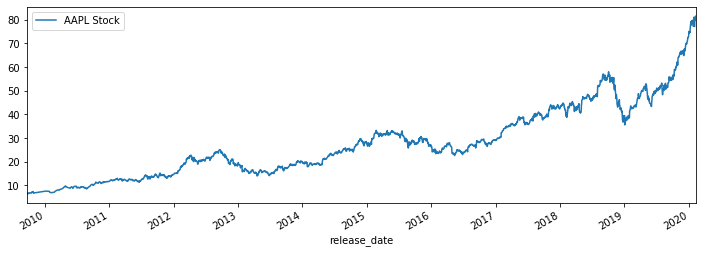
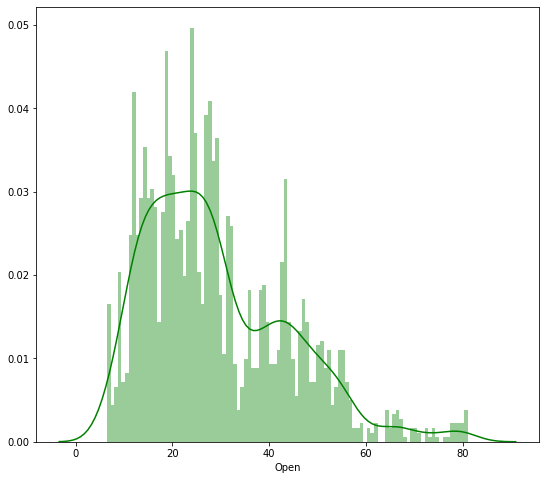
As financial institutions begin to embrace artificial intelligence, machine learning is increasingly utilized to help make trading decisions. Although there is an abundance of stock data for machine learning models to train on, a high noise to signal ratio and the multitude of factors that affect stock prices are among the several reasons that predicting the market difficult. At the same time, these models don’t need to reach high levels of accuracy because even 60% accuracy can deliver solid returns. One method for predicting stock prices is using a long short-term memory neural network (LSTM) for times series forecasting. An approach with the accuracy described above would be a very helpful toolset on hands of investors and especially in trading Companies. Knowing the next days stock price the investor/trader would be able to know where to put his/her money. Gradually with an accuracy more than 50% this business workflow would be profitable.

For solving the above business scenario we suggest the following methodology: we have used Neural Networks for predicting Apple and Microsoft Stock Prices, feeding our LSTM Network both with the sentiment of Financial News referring the stock and the previous days actual Open price. We suggest the following models to combat the above project:

* LSTM with Finbert Sentiment Analysis
* LSTM with Finbert Word Embeddings
* LSTM with stock prices and Sentiment Score Analysis.Finally, we compare the above implementation with ARIMA and make some predictions

**Data Sources**

Our data came from 2 data sources. The online news dataset containing titles and headlines about stock news is located publicly on Kaggle. The stock prices of the two companies we have analyzed (Apple and Microsoft) came from Yahoo Finance public API. Our Datasets of stock price seems to be imbalanced since stock prices of both companies are growing gradually the last years.



**Data Normalization**

Normalization is changing the values of numeric columns in the dataset to a common scale, which helps the performance of our model. To scale the training dataset we use Scikit-Learn’s MinMaxScaler with numbers between zero and one.

**Incorporating Timesteps Window Into Data**

We should input our data in the form of a 3D array to the LSTM model. First, we create data in 60 timesteps before using numpy to convert it into an array. Finally, we convert the data into a 3D array with X\_train samples, 60 timestamps, and one feature at each step. We are trying to predict the next day price, so the are look up step is 1 day.