Stock Movement Prediction Using Techniques of Deep Learning

Submitted to

American University of Armenia
Manoogian Simone College of Business and Economics

In partial fulfillment of the requirements for the degree of BA in Business

By: Arshak Mkhoyan
Supervisor: Dr. Vardan Baghdasaryan

Yerevan 2019

Abstract

A prevalent challenge in the field of portfolio management, stock movement prediction - through traditional means - has still not been brought to its optimum, in that there is yet not a ready-to-go algorithm for attaining guaranteed net positive returns from investment in stocks. This paper contests that, with the application of unconventional predictive methods from the realm of Deep Learning, it is quite possible to arrive at a satisfactory outcome in determining the direction of stock price changes. By constructing and training a model based on the 26-year long stock price data of a sizeable representative from a selected oil industry, 'Exxon Mobile', I arrive at a performance indicator of approximately 54%, which is significantly higher than a Random Walk benchmark of 50%. In consideration with the obtained results, this paper demonstrates that the designed predictive model is largely effective and, more importantly, successful in predicting the movement of stock prices.

Keywords: Stock Movement Prediction, Deep Learning, LSTM, Technical and Fundamental indicators.

Acknowledgment

I am greatly thankful to my academic supervisor Dr. Vardan Baghdasaryan for his suggestions, feedback and ongoing assistance. I also thank all the authors of materials that were used in preparation of this Capstone for their efforts and valuable insights. In addition, I would like to acknowledge the contributions of the alumni of American University of Armenia and my friend Yuri George Shakhnazarian for his continuous involvement in the project and demonstrated support. *The remaining errors, if present, are mine and only.*

Contents

List of Figures

Introduction

Stocks are regarded by many as one of the primary financial instruments that can tell the story about a particular corporation; a single stock represents a share of ownership in a company. Naturally, stocks have always been viewed as a very attractive investment by a variety of business agents; some even consider them a gamble, which can either make an individual's fortune or significantly drain their financial resources. What is so appealing about stocks is that they are easily accessible: stocks have been immeasurably popularized by the foundation of specialized stock markets (e.g. Nasdaq, NYSE), and, practically, anyone these days can get involved in purchasing and exchanging specific stocks, even if they do not possess any prior knowledge about stocks per se.

In particular, stock movement prediction has gradually become a heated topic of discussion in a business world. From likes of individual investors and brokers to mutual funds and stock issuing companies themselves, a number of players in the financial markets were and are presently interested in discovering effective stock price forecasting techniques. Since there is yet not a single, publicly available algorithm for a guaranteed success in predicting stock movement, investing in stocks remains very speculative, and, thus, concentrates such a passionate interest around it. Certain business sectors, particularly, are relatively untouched on such predictive methods, resulting in a great deal of uncertainties that lower investors' effectiveness in accumulating gains from changes in stock prices of companies from such sectors.

An example of one such 'unknown' sectors is oil. In spite of being very large monetarily and extremely important globally, oil industry has continually lacked a comprehensive assessment of stock movements of some of its major representatives. Though there are several instances of stock movement predictive attempts in oil industry, virtually none of them found wide recognition and application, mainly because such methods involved judgment by technical indicators, a judgment that deemed significantly inadequate in reality. While there is, clearly, a demand for a newer, functioning stock movement predictive tool in the oil sector, the supply for such a mechanism is just not there yet.

It is precisely one of the key reasons why I decided to draw my attention to designing a working stock movement predictive model for the oil industry. In the course of my academics, I have accumulated a substantial interest in stocks and especially determining the movements in their prices. Suited with necessary knowledge, experience and, after conducting a thorough research on the oil industry, expertise of the concerned sector, I was willing to investigate into unique solutions to current demands by applying innovative predictive algorithms in the scope of Deep Learning, an aspect of Artificial Intelligence and a branch of Machine Learning. For the purposes of my project, I considered focusing on the 26-year long stock movement of 'Exxon Mobile', one of the most sizable representatives of an oil industry. The primary objective of my experiment was set to achieving the satisfactory performance of the model, that is, arriving at a predictive model accuracy of a certain benchmark or more.

Throughout the paper, I will present a literature review on the valuable information found within external sources, a methodology on gathering and extracting the respective data used for training the model as well as the specifics of the model themselves, attained results and conclusive remarks.

Literature Review

In the course of my pre-experiment investigation, I have been continuously looking for valuable insights not only on the stock market and oil industry, but also on the predictive algorithms and technical specifications that could be tuned in the model. Throughout the conducted research, I got acquainted to a plethora of model variations both in and out of the scope of Deep Learning, which laid a bedrock for my future work. Furthermore, the analyzed papers served as a foundation to choosing right, accuracy-enhancing variables and refraining from the usage of less impactful ones. In this section I will present the key features and findings of distinguished works, which I continuously referred to during the project preparation and implementation.

One of the crucial, inspiring papers that I called a significant attention to was discussing stock price movements prediction using deep neural networks. In their work, Huynh Huy D., Dang Minh L. and Duong Duc (Huynh, 2017), outlined a comprehensive framework for assessing and predicting changes in stock prices by using a deep neural network algorithm called Recurrent Neutral Networks (RNN) and specifically delving into the application of Long-Short Term Memory Model (LSTM) and Gated Recurrent Unit (GRU). The authors of the paper suggest that most of the similar predictions were and even to this day are based on historical data, and thus seek to establish a different kind of model on different kind of input, incorporating other important predictors such as news in the meantime. In the conclusion, the experimenters assert that the performance of their optimized model was more than satisfactory, calling the model itself 'simple, but very effective' (p. 61). As a result, the paper provided a monumental evidence for the advantages that Deep Learning methods can generate in the realm of stock movement prediction.

A competitive edge to the practice of Deep Learning in stock movement prediction is also given by Charel Thiesen in his dissertation 'Predicting Stock Price Trends with News Headlines using Deep Neural Networks' (Thiesen, 2019). In the abstract, the author claims that 'The Efficient Market Hypothesis theory suggests that no model can predict stock price returns with an accuracy above 50% since the stock market is efficient and follows a random walk'. However, the results of Thiesen's work indicate that the predictive accuracy of his model averaged 60%, which, in accordance with judgments of a number of other scholars and experts in the field (Andrius Mudinas, 2018), is much greater than the benchmark accuracy of 50%. Abiding by the principles of Efficient Market Hypothesis, traditional predictive techniques cannot achieve a greater-than-50% accuracy, yet Thiesen's work once again demonstrates the predictive power of Deep Learning mechanisms.

On the other side of the coin, I was interested in constituting a catalogue of most effective, accuracy-boosting parameters that would have been used in my classification model. For that purpose, I referred to a variety of academic papers that encompassed such indicators. As a one example of such work is the publication 'Stock Price Movement Prediction from Financial News with Deep Learning and Knowledge Graph Embedding' by Yang Liu, Qingguo Zeng, Hunarui Yang, and Adrian Carrio (Liu, 2018). After thoroughly evaluating the authors' study, I concluded that the variable 'Stochastic oscillator (%K)' was of a massive positive contribution to the improvement of stock forecast, hence planning to include it in my model. Conversely, I decided not to revolve around news and similar sentiment processing indicators whatsoever**,** for I realized that the processing of news-related operations was somewhat out of my competence. Nevertheless, I do believe that news is a critically important aspect of stock movement prediction, and that is why I am looking forward to consolidating its adjusted variable component in my future models.

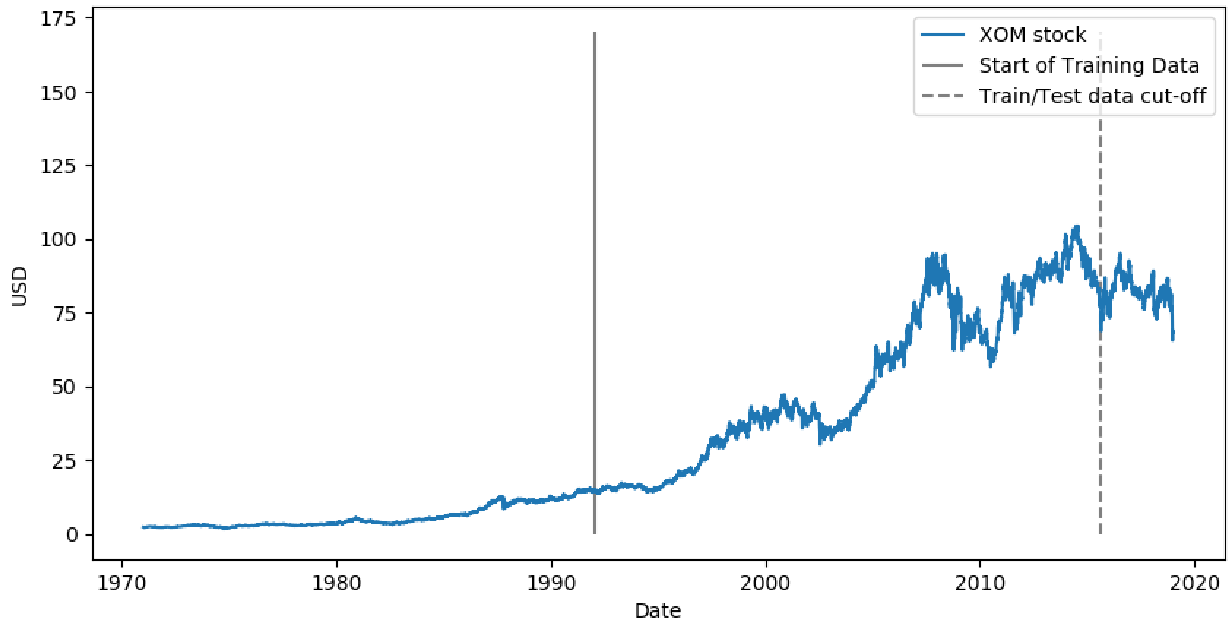
To sum up, based on the given insights from the literature about the factors and the models for stock price prediction, it has become apparent that forecasting stock price can be predicted more accurately than before. Although successful stock movement prediction can be difficult, is quite limited yet and contradicts the above stated Efficient Market Hypothesis, those challenges can be mitigated by the selection of appropriate Deep Learning instruments and inclusion of effective independent variables.

Data Description

As in all data-heavy projects, my first objective toward designing and implementing an effective stock movement predicting model was gathering an unbiased, in-depth data on stock price changes of the selected company from an oil industry, ‘Exxon Mobile’. It took me two weeks to complete the data gathering process, with all the necessary variables included there. The collected data was stored in a data frame of ‘Python 3.6’, an interpreted, high-level, general-purpose programming language. All the data preparation, feature selection, and model training and testing were performed in ‘Python 3.6’, too.

For the purposes of having a prolonged and comprehensive analysis, as well as fully reflecting major over-year differences in stock prices of ‘Exxon Mobile’, I was set to congregate a long-term data of at least 20 years of records. Eventually, I found a suitable dataset including 26 years of stock price changes, ranging from 31st December, 1992 to 31st December, 2018. Although I was able to find similar data of even longer timespan, I refrained from using it because in early years of ‘Exxon Mobile’ operation (1980s) there was little to no volatility (Refer to Figure 1) in stock prices of the corporation, thus gravely limiting the effectiveness of model’s predictability and deteriorating the overall accuracy.

***Figure 1**: XOM Stock Close Price*



The specified data was downloaded from ‘Yahoo’ financial website. The data extraction process was executed very conveniently, due to the fact that ‘Yahoo’ provides an in-built Application Programming Interface (API), via which it was possible to directly access the located data. It is also important to note that rows of the data, that is observations, showed daily records of stock prices. In addition, as is the general case for stock prices datasets, my dataset only included trading day information; in other words, it did not contain any records for weekend and holidays, therefore averaging around 253 daily observations in a single year.

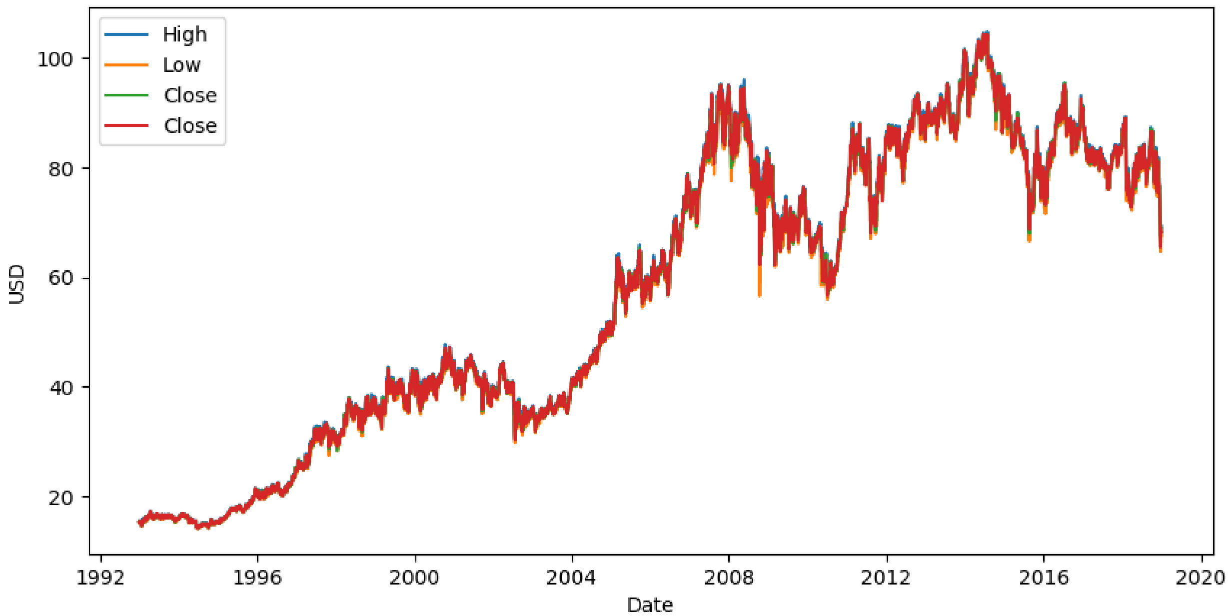
The transferred data included five columns (Refer to Figure 2), that is related components, for each of the stored observations:

1) ‘High’ - reflects company’s maximum stock price on a given trading day
2) ‘Low’ - reflects company’s minimum stock price on a given trading day
3) ‘Open’ - reflects company’s starting stock price on a given trading day

4) ‘Closed’ - reflects company’s final stock price on a given trading day

5) ‘Volume’ - reflects company’s number of shares or contracts traded on a given trading day

Figure 2*: XOM Stock Price Variants*



To quantify the accuracy of the model precisely, I decided to construct a target variable based on the initial components. Records for this target variable were obtained by subtracting the closing price of the company’s stock today from closing price of the company’s stock on the next day: if tomorrow’s price was higher than today’s, the target variable would be assigned a value of ‘1’, otherwise, ‘0’. Interestingly enough, in my dataset, specifically because it was structured on a daily basis, there were very few identical values of closing prices for any two subsequent observations. To accommodate for such an issue, I consulted with my advisor; based on his remarks, it became clear that the few identical values did not affect the model’s performance, therefore these could be assigned either ‘1’ or ‘0’, indifferently (eventually they were assigned ‘1’). Based on the derived target variable, it would then be possible to arrive at a sound estimate of the change in price direction and, therefore, productively gauge the expected stock price movement.

Among tools that can be massively helpful on the way of successfully determining the future price of a stock, experts signify two particular kinds of evaluation: technical and fundamental analyses. While, in principle, overly different from each other, both platforms provide a certain framework for properly assessing the estimated movement of a stock price and are by far regarded as must-know for any meticulous investor, speculator or, in my case, experimenter. In the next section, I will concentrate on the technical analysis, its aspects and components, more specifically delving into what technical parameters would be used in the scope of my project and what their impact would be. Thereafter, I will briefly talk about fundamental analysis, and conclude the topic of two analyses by synthesizing all the suggested variables within those analyses into a cohesive whole within my model.

Technical analysis

From the span of technical analysis, it is foremost that technical indicators are given a special attention. According to a 'stockstotrade' infographic, 'a technical indicator is a mathematical calculation that can be applied to a stock's past patterns, like price, volume, or, even to another technical indicator' (StockToTrade, 2017). Technical indicators are, as a matter of principle, different from such fundamental indicators as earning or profit margins in that they do not account for a stock issuing corporation's external/non-stock-related factors. Among the types of technical indicators, the four most generally recognizable are:

1) Trend indicators - relate to anything connected with trends of a stock price and patterns within them; such indicators gauge trends' direction, solidity and gravity

2) Momentum indicators - relate to an aspect of a change in stock prices associated with time; that is, insights and price ‘signals’ are generated based on differences in prices over time
3) Volatility indicators - relate to variability and unpredictability in stock prices; usually, these indicators deal with the rate of price movement and comparison between highest and lowest values of a stock price during some period of time
4) Volume indicators - relate to interrelation of stocks’ volume with the trend of their prices; volume indicators are concerned with measuring the volume-based impact on the strength of a stock price trend

Since each type of the technical indicators deals with relatively different characteristics of a stock-related component and is, therefore, fairly unique, I was inclined to integrate at least one instrument from each category in my model. As a result, and, as shown later, I was successfully able to do so, hence covering the majority of technical indicators and their aspects. In spite of this, I did not include as many instruments from each category as there initially have been, for throughout the empirical analysis I established a couple of interfering correlations between several such instruments, making my findings, overall, more biased. To avoid ending up with prejudiced results, I selected just a handful of, in my opinion, cornerstone and the most relevant technical variables, presented below.

From trend indicators, my choice of a to-be-included variable lied on an aggregate measure called ‘Moving Average Convergence Divergence’, or, MACD. The formula for MACD is as follows:

$$MACD = EMA_{12}(Close) - EMA_{26}(Close)$$

where $EMA_{12}(Close)$ stands for 12-period exponential moving average of close price, and $EMA_{26}(Close)$ stands for 26-period exponential moving average of close price. Itself, MACD is a time-series oriented technique that demonstrates and analyzes the relationship between two moving averages of a stock's price (Hayes, Moving Average Convergence Divergence, 2019). This particular instrument was selected mainly due to the fact that, in inference, the accuracy of my predictive model when MACD was utilized was substantially higher than when it was not. Also, methods within MACD are directly associated with time-concentrated data, which, in my case, could thus have been applied to the model very conveniently.

From momentum indicators, I decided to include Stochastic Oscillator, or, %K as their sole representative in my model. Inspired by the work of Huynh Huy D., Dang Minh L. and Duong Duc (Huynh, 2017), as mentioned in the section of Literature Review, I was willing to allocate some consideration for the instrument and, after thoroughly researching about it, was convinced that it would be a right choice to contain it within my variables. The generic formula of Stochastic Oscillator goes as follows:

$$\%K = \frac{(Close - L_{14})}{(H_{14} - L_{14})}$$

where *Close* is a close price of the given day, L_{14} is a lowest price in past 14 days, and H_{14} is a highest price in past 14 days. In accordance with Investopedia, 'a stochastic oscillator is a momentum indicator comparing a particular closing price of a security to a range of its prices over a certain period of time' (Hayes, Stochastic Oscillator Definition, 2019). In particular, bearing in mind the target variable that I have defined previously, it makes sense to take account of Stochastic Oscillator in my model. Not surprisingly, employing the instrument was found to be in good alignment with my goal of arriving at a higher predictive accuracy.

From volatility indicators, I referred to Average True Range as a volatility gauging measure. By definition, ATR deals with decomposing the entire range of an asset price for that period (Hayes, Average True Range - ATR Definition, 2019). Its formula goes by as follows:

$$ATR = EMA(max(High, Close_{previous}) - min(Low, Close_{previous}))_{14}$$

where EMA_{14}- 14-period exponential moving average.

Finally, from volume indicators, there are, in fact, three instruments that deemed worthwhile of inclusion. First, I chose 'Ease of Movement', or, EoM as a rudimentary attribute of relationship between stock's price and the number of issuances (volume). EoM opens up in a number of united interpretations and follow-up formulas as:

$$Distance\ moved = (\frac{(High + Low)}{2} - \frac{Prior\ High + Prior\ Low}{2})$$

$$Box\ Ratio\ =\ (\frac{Volume}{100{,}000})/(High - Low)$$

$$1 - period\ EoM = \frac{Distance\ moved}{Box\ Ratio}$$

$$14 - period\ EoM = MA_{14}(1 - period\ EoM)$$

where $MA_{14} - 14 - period\ moving\ average$ was found to be of biggest use as a variable, given my data. Secondly, to complement for EoM I selected 'Force Index', or, FI, which is 'an oscillator that measures the force, or power, of bulls behind particular market rallies and of bears behind every decline' (Chen, Ease Of Movement., 2018). Articulated as:

$$FI = (Close - Prior\ Close) * Volume$$

FI gives a good representation about the impact of a corporation's volume of issued stocks on those stocks' market prices.

Lastly, I decided to employ The Money Flow Index (MFI) as a different kind of volume-related technical component. In concordance with Stockcharts' description of a tool, 'MFI is an oscillator that uses both price and volume to measure buying and selling pressure' (Money Flow Index (MFI), 2018). Typically, MFI initiates from starting price for each period, and overall money flow is positive when there is buying pressure and is negative when the prices decline and sales are prevalent. Although quite complicated, MFI's formula is very helpful in understanding the relation between buying/selling pressure, total stock volume and current prices. That formula is presented below:

$$PP = mean(High, Low, Close)$$

$$PositiveMF = ((PP - PP_{previous}) > 0) * PP * Volume$$

$$TotalMF = PP * \text{Volume}$$

$$MFration = PositiveMF / TotalMF$$

$$MFI = mean(MFratio_{14})$$

Together with EoM and FI, MFI constitutes a set of techniques to effectively translate the volume insights into intelligence on future stock prices.

Learning about the kinds of and differences between technical indicators and understanding which of the instruments within these indicators would fit in my model, in hindsight, were crucial to the observed raise in test accuracy later on. Having technical variables set, I then aimed to respectively account for their fundamental counterparts; in the next section, thus, I will present my judgment on fundamental analysis and suitable indicators and instruments from there (For information on correlation among technical variables refer to Figure 5).

Fundamental analysis

To account for all the external, not stock-related influences on the price of a stock in due course, it is essential to consider fundamental analysis as an aiding tool in the process of stock movement prediction. In accordance with Investopedia’s broad definition of fundamental analysis, it ‘attempts to measure a security's intrinsic value by examining related economic and financial factors, which can be both qualitative and quantitative in nature’ (Ganti, 2019). Since the concentration of my research and predictions is for oil industry, it would thus make most sense to examine factors directly associated with the industry, prices of its primary products and national economies that in some way are related to it. Throughout the next paragraphs I will delve deeper into displaying and understanding the significance of this or that interacting attribute, pinpointing the specific metric for each and reasoning whether the chosen metric would be of a good fit to my model.

To form a better basis for choosing appropriate fundamental indicators I had to evaluate and, afterwards, select relevant national economies and key attributes within these that would have the largest direct impact on the development of and fluctuations in the oil market. After familiarizing myself enough with the topic, I have found ties from many external factors to such organizations as OPEC and countries as USA and Russia. While it might be intuitively clear how exactly are the specified actors associated with the oil industry per se, it is a relatively sophisticated task to determine applicable indicators linked to them. As a vivid example of such complexity, I could not find a usable, economic metric - on a day-to-day basis - for Russia, for anything else oil- and Russia-related, in terms of measuring dimensions, would be insufficient and of very little use for my purposes. Despite this shortcoming, I was able to gather several fundamental features that were in alignment with other mentioned, most important oil ‘stakeholders’ and that were taken into consideration upon design and exercise of the proposed model.

Starting with, perhaps, the most important country-contributor to the oil industry, and coincidentally the country of origin for the selected representative in ‘Exxon Mobile’ too, I had to ensure that U.S. economy is accounted for in my analysis. An instrument for that purpose was decided to be a U.S. treasury bond (with a maturity of 10 years) rate. While it is not uncommon to observe a U.S. bond serving as a representative of movements within U.S. economy, it made even greater sense to include that instrument in the scope of this project since USA is one of the major players in oil market (Ahmed, 2016) and likelihood would be quite high that if, say, one of U.S. major, oil industry-based corporations flourishes, US bonds’ perceived safety would increase and their interest rate, as a result, would drop. As a confirmation of the adverted hypothesis, I found that the rate of U.S. treasury bond correlates negatively with stock price of ‘Exxon Mobile’ in approximately 86% (Refer to Figure 6). However, the implication of this finding was that including an indicator for U.S. treasury bond’ historical data in my model would be unjustified from the principle of avoiding biased outcomes. Hence, although the effect of U.S. economy was generalized and reflected within context of stock movement prediction, the application of this particular fundamental instrument was abandoned.

Thereafter, I looked for suited historical data on OPEC. More precisely, I contemplated focusing on corresponding individual economies of major OPEC delegates: out of such delegates I prioritized the assessment of Saudi, Irani and Iraqi influence in oil industry. In spite of this idea, throughout a number of references (Islamic Republic of Iran, 2018), (Saudi Arabia: Economy, 2017) it caught me clear that all three economies are extremely oil-reliant, and, instead of opting for some concrete measurements for each of the countries, I decided to incorporate world oil price as a joint fundamental instrument for the evaluation of OPEC's significance to the realm. As expected, world oil price was found to be of gargantuan association with stock price of 'Exxon Mobile', where increase in one indicator would usually result in a closely equivalent increase in other (Refer to Figure 6). Again, as was the case with U.S. treasury bond, I would abstain from using world oil price as an independent variable in my model, for consequent bias and overfitting from its inclusion would be tremendous.

Continuing on the discussion of world oil price, it was crucial to recognize that because oil is traded globally in U.S. dollars (Farley, 2019), the U.S. currency itself could act as an important external factor to the movement of stock prices in oil industry. For the provided reason, I intended to perform correlation check of 'Exxon Mobile' stock price with U.S. Dollar Index - DXY. DXY, according to Investopedia, 'is a measure of the value of the U.S. dollar relative to the value of a basket of currencies of the majority of the U.S.'s most significant trading partners; [DXY] is similar to other trade-weighted indexes, which also use the exchange rates from the same major currencies.' (Chen, U.S. Dollar Index - USDX Definition, 2019). While DXY's correlation with the element of our interest was not very significant, rounding -46% (Refer to Figure 6), fusing U.S. Dollar Index into my model, again, proved lackluster as no boosts in performance accuracy were registered.

Finally, I decided to switch my attention from national economies and oil-related metrics out of there to market-oriented practices and indexes. Notably, I examined the relation between stock price of 'Exxon Mobile' and market strength and market volatility measuring indices 'S&P500' and 'VIX', respectively; while both indexes are not particularly addressing oil industry per se, they are still maintained as great indicators for general assessment of a certain industry. Though the two attributes showed fairly little (about +46% correlation in case of 'S&P500') to no (approximately zero correlation in case of 'VIX') association with stock price of 'Exxon Mobile' (Refer to Figure 6), none of them improved the model so much as to be considered for the inclusion in its final version.

In summary, every single one of the proposed fundamental instruments, be them distinct indicators with respect to selected national economies or market-capturing gauges, failed to enhance my model's performance to a point in which overfitting and excess bias would be avoided, whereas accuracy would be optimized. As a result, technical and fundamental components were not synthesized into the model as a one or several denotative variables; instead, only technical instruments found their use. In the next section, I will present the introduction to the constructed model, the model's detailed specifications, graphical representation and test outcome. Then, I will conclude the paper by illustrating key findings of the experiment and emphasizing future work to be done.

Model Architecture

Having all the preparatory steps of the project emphasized in detail, I now shift to providing the description of my model and its key elements.

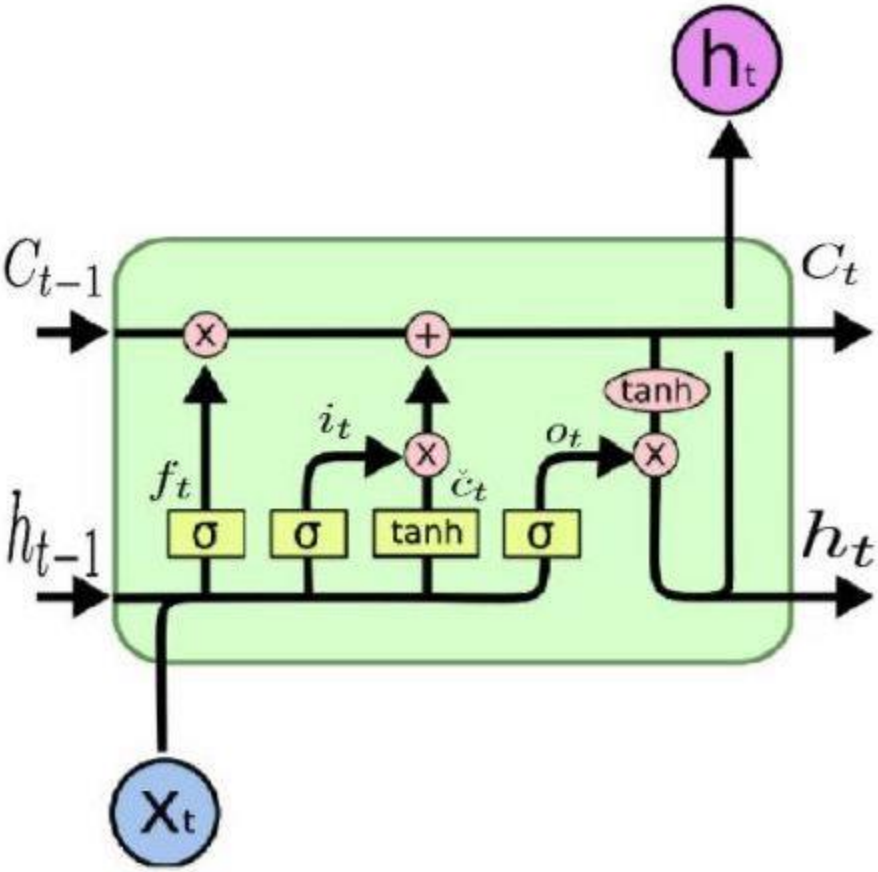
As mentioned previously, I decided to design a stock movement prediction algorithm by applying conceptual theory from the realm of Deep Learning. Supported by a number of academic materials that were discussed in the Literature Review, Deep Learning techniques - when used properly - were found to be significantly more effective in achieving satisfactory model accuracy than conventional methods of prediction. In particular, for the purposes of my experiment I referred to the application of Recurrent Neural Networks (RNN), since RNN is used notably for sequence data, one type of which is the time series data that I had already gathered.

Because RNN contains a vast quantity of miscellaneous models, I narrowed the scope of my analysis to one specific model called Long Short-Term Memory (LSTM) unit. Why LSTM was prioritized over other attractive models from family of RNN is due to the fact that it was eventually found superior to all of them, and, hence, I focused on construction and development of it only.

Variables ‘Close price’, ‘Trading volume’ and all those corresponding to technical indicators prescribed previously served as a final, colligated input to my LSTM model. As for variables corresponding to technical indicators, their ‘time step’, that is, how much of the previous data, in terms of its time span, were to be assessed for the selected features for the price movement prediction of the next observation, was set at 5 working days. Essentially, each observation in the input data with its respective variables was in a form of 5x9 matrix, in which 5 rows were consistent with the selected time step and 9 columns were consistent with the number of included features (‘Close price’, ‘Trading volume’, and 7 variables corresponding to technical indicators). As a final note to technical variables, I need to underscore that these were accordingly adjusted for standard deviation before actually being processed in the model. Additionally, input data was also scaled using ‘MinMax’ method.

The canvas of my model is portrayed below. The formulae for each of the model’s explicit components are afterwards pinpointed, followed by the thorough description of how the model is supposed to operate and how it arrives at the desired outputs, which are then used for accuracy evaluation.

Figure 3: LSTM Cell



The key components of the LSTM unit are presented below. Note that $*$ in formulae acts as elementwise multiplication. In addition, for best understanding of what the concepts stand for, it is recommended that a reader get acquainted with the general description of the model first.

- Forget gate (f_t), which is given by

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

This gate judges whether the loaded information should be kept or thrown away. The closer the received values are to 0, the more this gate is inclined to 'forget' the information, while the closer the received values are to 1, the more this gate is inclined to 'memorize' and subsequently store the information.

- Input gate (i_t), which is given by

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\check{C}_t = tanh(W_C[h_{t-1}, x_t] + b_C)$$

Input gate is there to regulate the current Cell state (presented next). The gate receives previous hidden state and current input and exposes it to sigmoid and *tanh* functions, with the help of which it updates the current memory of the model.

- Cell state (C_t), which is given by

$$C_t = f_t * C_{t-1} + i_t * \check{C}_t$$

Having gone through forget and input gates, it is possible to derive the current cell state. First, the previous cell state gets pointwise multiplied by the forget vector. Then, the output from the input gate is added pointwise to the current information, after which the new cell state is formed.

- Output gate (o_t), which is given by

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

After certain modifications are done within input gate, the new hidden state is transferred to the output gate. Essentially, the output gate decides what the next hidden state should be.

- Hidden state (h_t), which represents output and is given by

$$h_t = tanh(C_t) * o_t$$

Hidden state represents the final output of one cycle of the model. In its essence, it is obtained by multiplying values of output gate to the current cell state. The combination of all hidden states then provides the final output unit.

Before proceeding to the description of the model itself, it is important to stress that input, output and forget gates have the exact same equations; the difference within them, however, comes from the difference of weights assigned to each gate, as defined by the model specification.

In my model, each observation is assigned its unique vector. As is shown in the canvas above, a vector goes through input, output and forget gates, independently from each other. Once a vector passes through these gates, the sigmoid activation function, represented by σ , and defined as follows:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

outputs values between 0 and 1, which after being multiplied elementwise with similar values of another vector, determine the proportion of that another vector that is "let through"; say, if another vector is close to generalized value of '1', the majority of that vector is 'stored' in the memory, and 'emptied' if vice versa.

Whereas the input gate states how much of the newly computed state for the current input should be let through, the forget gate states how much of the *previously* computed state should be done so. The received vector then is processed through hyperbolic tangent activation function, represented by $tanh$, and defined as follows:

$$tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

that scales the data from -1 to 1. Finally, the output gate then defines what should be transferred further from the updated memory. The outlined process is then repeated as many times as is the length of a time step.

Eventually, the model arrives at a 'hidden state', h_t, which rules how the previous memory and the new input should be combined together. Since the final hidden state should have the same number of units as is the size of dimensions for all three gates - chosen by a model architect as a hyperparameter -, h_t contained 13 elements, in accordance with the number of dimensions that I selected for my gates. The 13 units are then connected to a final output unit, which is defined in the following way:

$$\hat{y}_t = \sigma(Wh + b)$$

where W is the weight array, h is final output from LSTM, and b is bias term.

After receiving the final output, the obtained values are compared to actual ones via the cross-entropy loss function, which is minimized for the losses between actual and predicted values by using backpropagation update weights and is given as follows:

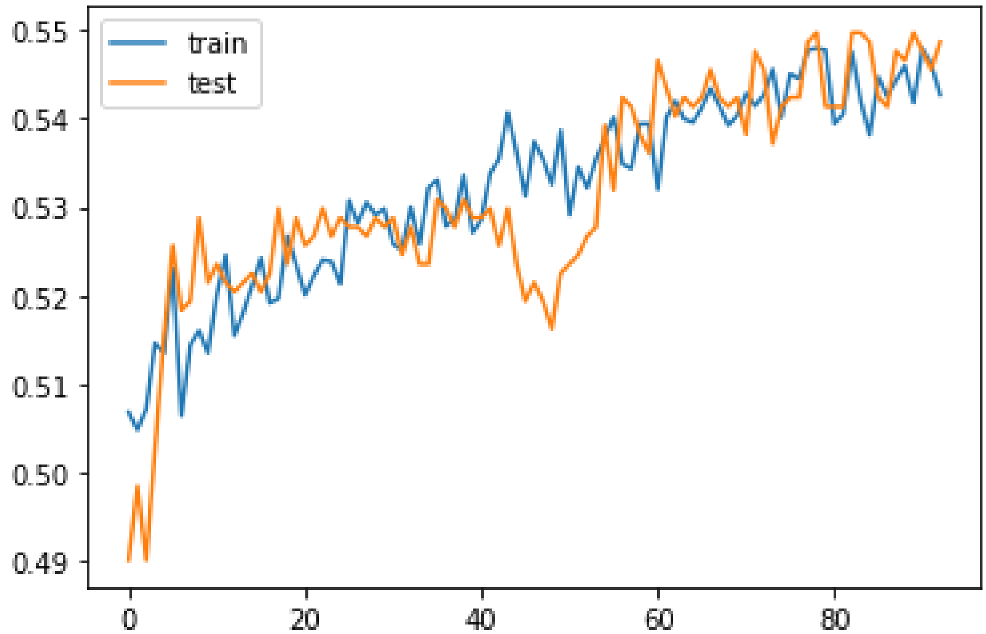
$$L(\mathrm{y}, \hat{y}) = \sum_{1}^{n} -(y * log(\hat{y}) + (1 - y) * log(1 - \hat{y}))$$

where n is the number of observations in training data, y is actual values of target variable and $\hat{y}$ is predicted values. In the ultimate end of the operation it is already possible to arrive at the performance accuracy of the model and attain measurable results (For more information with regards to how LSTM models work in general, a reader is welcome to refer to the following article).

Results

The results that I achieved after running the finalized algorithm can be summarized in and interpreted from the accuracy graph below:

Figure 4*: Accuracy of LSTM model, train/test*



Although quite volatile, the accuracy line for test-allocated data peaks around 54%, at the number of train cycle epochs of about 90. Though in the early stages of training the accuracy is only around 49-50%, just as could be expected in abidance with Efficient Market Hypothesis, the more does the model cycle through the train-allocated data, the more accurate does the prediction become. Interestingly enough, there was no significant increase in accuracy registered when the model would cycle over more than 90 epochs, for overfitting would deteriorate the results to a point where model's performance would actually be worsened rather than improved.

In conclusion, the LSTM unit's accuracy is, in general terms, quite satisfactory, thus verifying the hypothesis that it is feasible to have an upper hand in predicting the movement of stock prices.

Conclusion

With all the preparatory steps, including but not limiting to topic and industry research, literature review, analysis of technical and fundamental indicators and model design, execution, and assessment, properly undertaken, I have arrived at several important findings within this experiment and, most importantly, an answer to my initial research question: whether Deep Learning techniques can be effective in predicting stock price movement.

This work illustrates that, in the frames of oil industry, it is indeed possible to 'beat the market' by a margin of approximately 4% by applying unconventional, non-standard predictive methods. Of other considerable takeaways from my project that might be helpful to a similar research is a revelation that the most impactful predictors were the past stock prices, trading volumes, and technical variables, while the least impactful ones were among fundamental indicators.

In future, I am planning to perfect this project by refining my working model, accommodating for the factor of world news and investigating on the subject matter even more comprehensively, and, perhaps, to broaden the spectrum of the conducted work to other business industries, using potentially more enhanced or newly discovered Deep Learning practices. Having come at a satisfactory outcome within the scope of this experiment, I am ready to get involved into any kind of alike study, in the course of which I will, again, suggest and religiously aim to discover the benefits that Deep Learning can provide, benefits deemed generally unrealistic before.

Appendix

__Figure 5__: Correlation Matrix of Technical Indicators

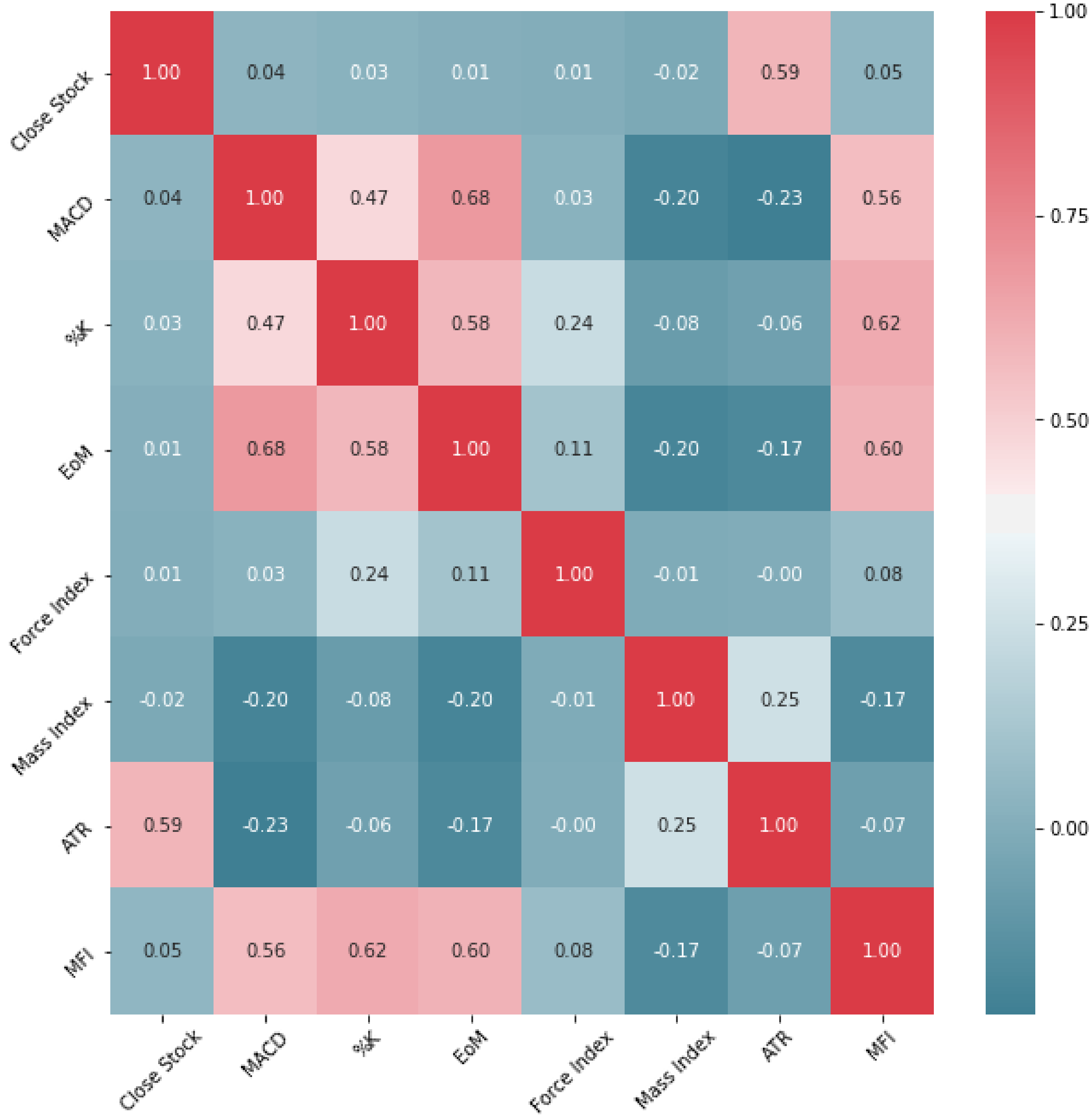
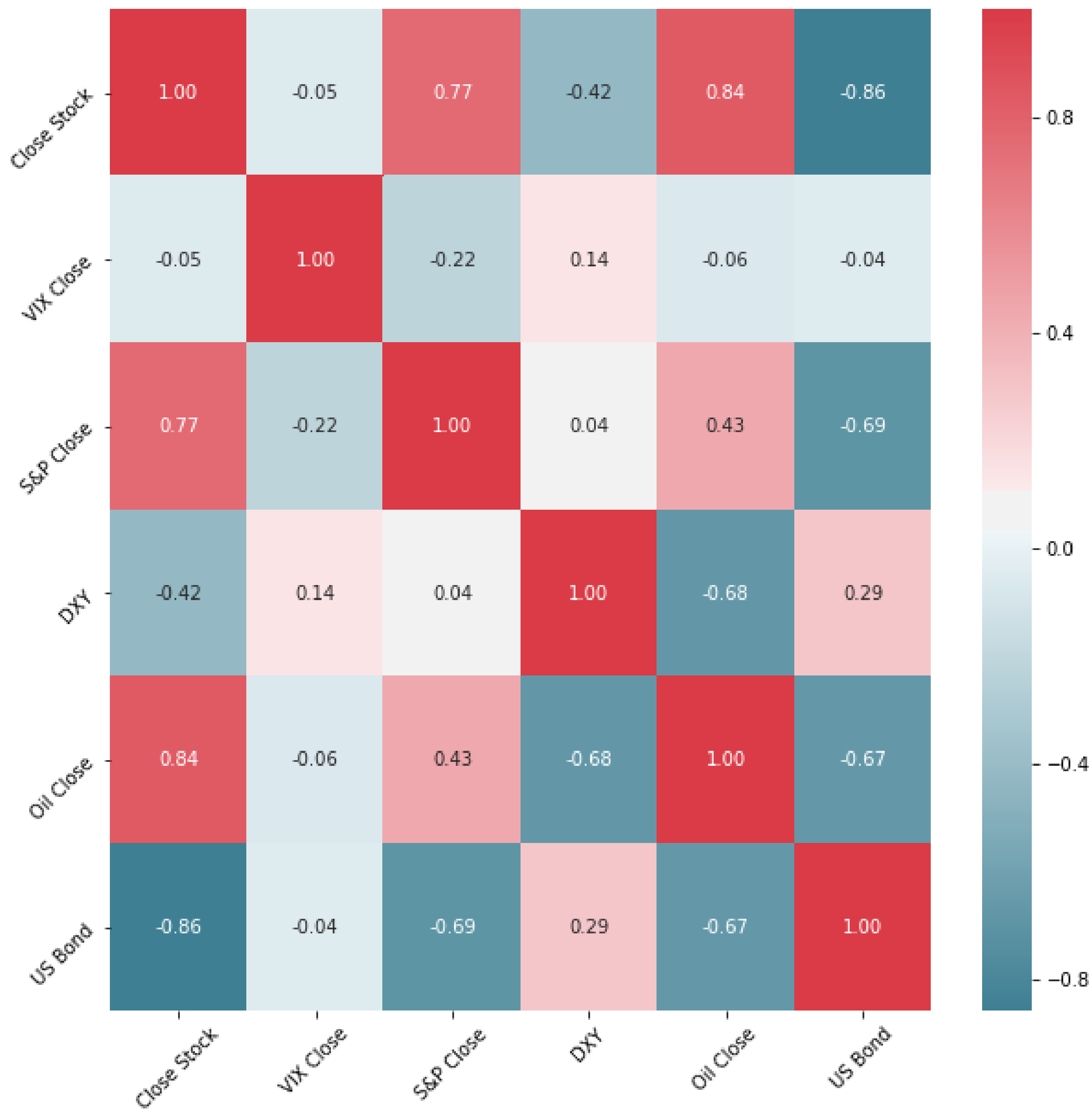


***Figure 6**: Correlation Matrix of Fundamental Indicators*



References

Ahmed, S. M. (2016). *The Impact of Oil Prices on the Economic Growth and Development in the MENA countries*. Middle East Economic Association.

Andrius Mudinas, D. Z. (2018). Market Trend Pre- diction using Sentiment Analysis: Lessons Learned and Paths Forward. *The 7th KDD Workshop on Issues of Sentiment Discovery and Opinion Mining*, (p. 10). New York, NY, USA.

Chen, J. (2018, May 15). *Ease Of Movement.* Retrieved from Investopedia: https://www.investopedia.com/terms/e/easeofmovement.asp

Chen, J. (2019, April 10). *U.S. Dollar Index - USDX Definition*. Retrieved from Investopedia: https://www.investopedia.com/terms/u/usdx.asp

Farley, A. (2019, May 19). *Oil & Currencies: Understanding Their Correlation*. Retrieved from Investopedia: https://www.investopedia.com/articles/forex/092415/oil-currencies-understanding-their-correlation.asp

Ganti, A. (2019, April 1). *Fundamental Analysis*. Retrieved from Investopedia: https://www.investopedia.com/terms/f/fundamentalanalysis.asp

Hayes, A. (2019, May 5). *Average True Range - ATR Definition*. Retrieved from Investopedia: https://www.investopedia.com/terms/a/atr.asp

Hayes, A. (2019, May 9). *Moving Average Convergence Divergence*. Retrieved from Investopedia: https://www.investopedia.com/terms/m/macd.asp

Hayes, A. (2019, May 8). *Stochastic Oscillator Definition*. Retrieved from Investopedia: https://www.investopedia.com/terms/s/stochasticoscillator.asp

Huynh, H. D. (2017). A New Model for Stock Price Movements Prediction Using Deep Neural Network. *Proceedings of the Eighth International Symposium on Information and Communication Technology - SoICT 2017. doi:10.1145/3155133.3155202*.

Islamic Republic of Iran. (2018, October 11). Retrieved from The World Bank: https://www.worldbank.org/en/country/iran/overview

Liu, Y. Z. (2018). Stock Price Movement Prediction from Financial News with Deep Learning and Knowledge Graph Embedding. *Knowledge Management and Acquisition for Intelligent Systems Lecture Notes in Computer Science*, 102-113.

Money Flow Index (MFI). (2018). Retrieved from StockCharts: https://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:money_flow_index_mfi

Saudi Arabia: Economy. (2017). Retrieved from Global Edge: https://globaledge.msu.edu/countries/saudi-arabia/economy

StockToTrade. (2017, April 25). *Types of Technical Indicators*. Retrieved from StockToTrade: https://stockstotrade.com/types-technical-indicators-infographic/

Thiesen, C. (2019). *Predicting Stock Price Trends with News Headlines using Deep Neural Networks*. London: City, University of London.