# How Deep Learning Can Be Used in Stock Price Prediction Research Paper

## 1. Introduction

Current stock price prediction methods have received negative attention for the variety of assumptions placed on the market. The misuse of these assumptions have misguided the predictions made by experts and researchers and have proven that limiting the variables within these methods does influence the change in trading signals when making predictions. Today, technological statistical advancements and system analysis are used most commonly to extract mass amounts of data on the stock market in ways that human capacity and methods cannot. Major news websites are giving information out to the public (on all types of media). For example the amount of consumers that visit home page of a major internet search provider when they want to view the stock market. Also to extend on this as technological advancements grow in sophistication their uses can grow too. We now have the ability to store information digitally (where the limit is that of our capability to store information on a physical medium it also grants us the ability to manipulate information too. Social media analysis differs to typical methods as they use information from a wide range of sources. This provides us with a format to discover information that previous methods could not.

### 1.1. Background

Stock price prediction has been a topic of interest for economists and analysts for a long time. Traditional methods mostly included technical analysis and stock predictions which have had limited success. With the flourishing of artificial intelligence and machine learning, the predictive power has gained attention because of the perception that it is now possible to predict based on historical data. This has brought about better and significantly improved stock market predictions.

### 1.2. Problem Statement

The most challenging task in finance is to predict stock prices, and the stock price is considered to be non-linear and complex time series data, so it is difficult to identify its pattern . This is where the role of deep learning comes into picture. To predict stock prices, many financial companies and investors have been developing and using different models that can have effectively used.

### 1.3. Objectives

Identify and use an optimal, efficient, and faster strategy to forecast the stock market value and predict its behaviour using deep learning models. Discover and evaluate potential improvements in current strategies in order to achieve the goals aforementioned, effectively understand and apply various relevant and applicable concepts of the subject of artificial intelligence and pattern recognition, and apply all of them to the task of stock market forecasting with the hopes of improving the probability of achieving results that are beyond the current, state-of-the-art and not yet achieved by any other team.

## 2. Literature Review

In literature it is hard to find a detailed description of the problem applied to the stock world, as well articles and CHI Papers seem to be the main resource for some undetailed description of the predictive techniques. Regarding the practical part of the thesis, primary quantitative data were used by collecting 4 years of stock shortage on IBM, Google and Amazon during 2013-2017, and using 1-year forecast algorithms. After this period, the predictive models will be implemented and the results will also be compared and evaluated in order to find out the accuracy of the initializing prognosis.

### 2.1. Overview of Stock Price Prediction

The stock market plays a vital role in the growth and development of the economy (Ryan, 2013). Predicting the future price of stocks is a very difficult and complex task. Stock price prediction is the act of trying to determine the future value of a company stock and season investors are always looking for new methods to form the predictions. The most clear reason for this is to take advantage of the expertise. If investors can predict a stock’s value, they can take advantage of creating money. There are many methods of predicting stock outcomes, but this paper will focus on how deep learning can be used in stock price prediction.

### 2.2. Traditional Methods in Stock Price Prediction

As opposed to the review of the most state-of-the art knowledge, indicating on the likely diversities between convention techniques and deep learning has been employed in Stock Price Prediction may generate a more valuable as well as attractive topic. Furthermore, traditional methods in stock price prediction may help strengthen a basic grasp on financial trading, allowing to progress to this particularly complex convolutional neural network and long short-term memory neural networks that are deep learning methods. Traditional methods could be utilised as a benchmark to see how data may be researched and improved through deep learning in Stock Price Prediction, enhancing our understanding of the overall topic.

### 2.3. Introduction to Deep Learning

Safeguards like the 2005 Supervision of Single Stock Futures Act that require reporting of large trading positions, have been put in place in order to prevent insider trading. These efforts to create an equitable market economy and transform present ways to comply with modern day issues are difficult and currently ineffective. New, innovative ways are necessary to counteract the inadequacies of present methods.

## 3. Methodology

We then proceed to select suitable deep learning models that will be used to predict stock prices in this research. In the selection process, profitability of the model, unit root test and final validation accuracy score according to r2 score will be used to assess the importance of a model. The deep learning models to be selected will be Long short term memory (LSTM), Gated recurrent unit (GRU) and simple feed forward neural network. Long short term memory will be used to predict stock price during non-volatile periods, GRU during volatile periods and the simple feed forward neural network will be used as a benchmark model.

### 3.1. Data Collection and Preprocessing

The time series data of stock price for the model is collected from the official Yahoo fiance webpage. The perhaps models are trained and tested on the basis of two metadata namely the Open Price and the Close Price of any specific stock. The sets used for training were from the period of January 2011 until September 2016 in the same manner as the testing set was from the period October 2016 to January 2017.

### 3.2. Deep Learning Models Selection

3.2. Deep Learning Models Selection: One of the major problems with the stock price prediction is how to effectively model the time series data. Deep learning appears promising due to its capability of accurately handling the complex and nonlinear relationships in data. Therefore, we would like to compare the performance of different kinds of deep learning model on this task. First of all, we benchmark the deep learning models to the traditional time series modeling techniques. This is a critical step as we aim to confirm whether the data indeed shows nontrivial patterns that can be exploited by deep models. For instance, some bad news was assumed to have larger impact on stock price compared with good news. After the environment is confirmed, we then construct and compare various deep learning architectures, such as stochastic recurrent networks (SRN), long-short term memory (LSTM) networks, and time-induced neural networks (TNN).

### 3.3. Training and Testing Process

The data have to be separated into two parts, the training subset and the testing subset. The training subset is utilized to build the training model and the testing subset is utilized to evaluate the precision of that model to predict. In contrast to traditional econometric models, where a common practice is to splice the data in chronological order, to generate inputs and outputs, it is necessary to use a sliding window method to repeatedly train and test the accuracy of the model. The correct approach to train the model is to select a starting point and use all the data before that point as the training subset, for example, if the data goes from 2006–2015 the first time I will use the data from 2006–2007 to train and 2008 to test it. Then I would add to the training subset all the data values from 2006 to 2008 and I will use 2009 to test it. The model will only be trained with the data subset it has already tested so far. It is important to mention that the reason to walk the model in that way is to be sure of avoiding possible data leakage, once the investor would not have known the future data.

## 4. Deep Learning Techniques for Stock Price Prediction

The second deep learning technique is the Recurrent Neural Networks (RNN), a class of artificial neural networks designed to recognize patterns in sequences of data and have been successful in predicting sequential data. More specifically, an RNN was designed to include cycles that allow data to persist, hence the most popular application is text predictability. An RNN has two inputs, the sequence and the state at time t-1, which generates the state at time t. Hence, RNNs consider past information while predicting the future. Nonetheless, RNNs are limited to implementing the recursive cycles, which requires significant data to accurately predict stock prices.

### 4.1. Convolutional Neural Networks (CNN)

In Convolutional Neural Networks (CNN), the data is usually made up of grid structures such as images, though time series data can be converted into an image-like structure. The CNN consists of smaller units called neurons and several layers. The three types of layers are convolutional layers, pooling layers and fully connected layers. The convolutional layers are made of filters of specified sizes which move about datasets generating several versions (feature maps) and forming feature hierarchies. The pooling layers group together neurons of previous levels into single neurons. These two types of layers give the model inbuilt translation invariance. The fully connected layers operate on the final feature map to form the final predictions. Moreover, CNN frameworks are processing powers although not as fast as other models because more parameters will have to be estimated. However, samples required for training are fewer. Pick data costs very highly on time.

### 4.2. Recurrent Neural Networks (RNN)

RNN also consists of a series of nodes that transmit the output to the next node in the sequence. The output from the previous node is treated as the input of the current node. Accordingly the hidden state from the previous node acts as the input of the present one. In stock price prediction, the parameters pertain to historical stock prices instead of characteristics, which are used for speech and language processing. In order to predict stock prices for any given day, the historical prices for a number of days before the given day are used as input to serve as the dependent variable. The independent variable is the predicted stock price for the next day. Predictions are granted for stocks that are publicly sold without interference from a company or an individual. These stocks are also accessible for public education and do not have any insured or private information affiliated with them.

### 4.3. Long Short-Term Memory (LSTM) Networks

LSTMs are the preferred algorithms for deep learning model as they can help to tackle the vanishing or exploding gradient issues experienced in vanilla recurrent neural networks. This is why they are especially suitable for time series prediction tasks. When the time step t is on the smaller side, one can potentially use a classical and type of RNN cells and add a dense layer on top of them to turn a quote on quote ''vanilla'' RNN into a forecasting model. However, when the sequence gets longer, LSTMs could have their own recurrent network which can handle long term dependencies, where one has to look far back in the past to do the predictions. This is illustrated in the following Figure.

### 4.4. Generative Adversarial Networks (GAN)

Generative Adversarial Networks (GAN) have the ability to produce extremely convincing synthetic data, specifically in the image and video domains. A GAN is actually composed of two separate networks, the generator, and the discriminator. The generator is responsible for producing data samples given a randomly weighted input vector, and the discriminator attempts to discriminate between real and synthetic data samples. The goal is for the synthetic data to eventually be so close to the real data that the discriminator can no longer differentiate between the two. Once this equilibrium is met, the GAN can be considered to have ‘learned’ the underlying data distribution. GANs have shown extreme promise in stock price prediction, as synthetic data can be used to run the generated inputs through an existing prediction model much faster, as it is not limited by the time required to download and process real data.

## 5. Evaluation and Results

Being able to obtain a predictive result, of a change in the stock price of a company, is one of the greatest advantages for an investor. Deep learning is a branch of the machine learning algorithms that have been shown to have excellent performance in several fields, especially natural language processing and image recognition. This paper will work on designing a Long Short-Term Memory Stock Price Predictor (LSTM-SPP) model that has outperformed several traditional models and machine learning algorithms. The major aim of this paper is to estimate the performances of the LSTM-SPP using several performance metrics and compare the obtained results with that of the traditional models such as Moving Average and Auto-Regressive Integrated Moving Average (ARIMA). However, due to length constraints of this paper, we present four out of several performance metrics, which are most commonly used by the researchers and they are Mean Squared Error (MSE), Mean Absolute Error (MAE), Symmetric Mean Absolute Percentage Error (SMAPE) and R Squared (R). A long time-series daily data of a giant company, ‘3M Co’, will be used in order to make sure the efficiency of the introduced model. For the analysis part, we are also using three driving stocks, ‘SPY’, ‘GE’, and ‘IBM’, to increase the volatility of the market. The structure of this paper is as follows; in the second section of the paper, which is the literature review, I will be presenting the related works on this issue. Publicly traded companies are required to produce a quarterly and an annual financial report. In addition to these reports, corporations often hold earning call meetings quarterly. These meetings and reports contain information about the speculation of the corporate’s progress, future plans, and many other useful data for the investors. However, the only problem is that the release dates of the reports and the meetings are spontaneous and unpredictable. This feature of the data is called the noisy data, and it is a common issue in financial datasets. There are several machine learning models have been developed in order to resolve the noise problem. Also, there exists two major kinds of models in stock price prediction literature, which are the traditional models and the modern models. The traditional models, like Moving Average, Moving Average Convergence Divergence, and ARIMA, are based on computerized results of the financial indicators in the past. However, the more modern models, like Support Vector Machines, Artificial Neural Networks, Long Short-Term Memory, and Wavelet Neural Networks; are based on the raw data. At this point, the capability of the modern models in using the apparent data is the main reason why the deep learning and artificial neural network researches have been focused more instead of the statistical or econometric researches.

### 5.1. Performance Metrics

When evaluating models, it is important to have metrics that allow for quantitative comparisons. Most common metrics used for models based on financial time series are accuracy, precision, recall, the F1 score, and the correlation coefficient. The accuracy measures the number of correct predictions made as a fraction of all predictions. The precision is the number of true positives as a fraction of true positives plus false positives. The recall is also known as the true positive rate or sensitivity, which is the number of true positives as a fraction of true positives plus false negatives. The F1 score is the harmonic mean of precision and recall, which allows for a balance between the two metrics. Finally, the correlation coefficient is a measure that describes the strength and direction of a linear relationship between two variables. For a back tester, which corresponds to a trading algorithm, the Sharpe ratio, the Sortino ratio, and the maximum drawdown are also generated. The Sharpe ratio measures the risk-adjusted returns, normalized by the standard deviation, the Sortino ratio measures the risk-adjusted returns, normalized by the downside risk, and the maximum drawdown measures the largest peak-to-valley loss. These performance metrics have been combined as a framework, which evaluates the financial time series predictions and the number of times the model correctly predicts the direction of the next closing price, and the false signals generated by the model.

### 5.2. Experimental Setup

Now we will explain how the tests were done and present the numbers. In order to do that, we must first explain which stock prices were chosen. We used prices of Apple, Inc, for the period from. Within that period there are 1504 trading days in total; we trained our networks for the first 754 days; we validated for the next 150 days; and we tested for the last 500 days. For the training and validation partitions, the first 600 trading days were used to create input-output pairs of six trading days of input and one day ahead output. To illustrate, the first training data example was formed with the prices from trading days; where the input vector was of length six was conformed by the share prices of \(P\_{1,1} = 21; P\_{1,2} = 21.19; P\_{1,3} = 20.95: P\_{1,4} = 20.66; P\_{1,5} = 21.18; P\_{1,6} = 21: and the output was 21.11.

### 5.3. Analysis of Results

However, as we have shown in this work, the latest advances and approaches in deep learning, like LSTMs and CNNs, give a good performance, because of the sequential nature of the data and the patterns these algorithms can capture. Many crime incidents require solutions and responses in real-time. In an ideal scenario, predictive policing aims to predict where the crimes are likely to occur and send the police force there in advance. For such purposes, the model needs to be trained on real-time data which is huge in terms of record length. It comes from various sources like video data, age of the criminal, and various other stage demographics.

## 6. Discussion

Scientists can't seem to agree on whether AI will save humanity or destroy it. And what's fascinating is that these conflicting predictions often come from within the ranks of AI researchers. I consider myself a technological optimist - I believe that AI will keep improving our lives in ways which we can hardly imagine yet. But I often find myself debating with other AI researchers who take a more pessimistic view of the future. From my observations of these debates, the bigger 'yes or no' questions about AI's impact on the world have now splintered into countless sub-controversies which are equally thorny. This is one reason why public perception of AI seems to be in a state of constant turmoil. Every day seems to bring new stories about AI curing cancer and AI stealing jobs, often from the same news source!

### 6.1. Comparison of Deep Learning Techniques

We examine four different convolutional neural network models that are known to be efficient in stock price prediction. These networks are Alexnet, GoogleNet v1, VGG16 and ResNet 50. For each of them, we describe the input layer, the convolutional layers, the fully connected layers and the output layer. We also compare the predictive performance of these networks using 2 different stock indices and using different time periods for these indices.

### 6.2. Limitations and Challenges

First, when applying deep learning models to stock price prediction, having large quantities of data is necessary, and collecting this data is not always easy. Additionally, like traditional time-series models, the model could not work well when the stock price is affected by rare unpredictable events like war and natural disasters. Second, in reality, it is controversial whether the capital market is random walk and does not follow any patterns or not. The efficient market hypothesis suggests that the current stock price fully reflects all available information and the factors affecting the stock price in the future can be unpredictable. It is not known to what extent deep learning models can reflect all available information and extract patterns in the data since the market is affected by countless factors. Third, the deep learning model demand large computational power and resources or the employ of several layers of network and this could let to high electricity cost and to a high use of technology devices, in consequence smartphones which manufacture produces an environmental impact and to a slow network signal. Also, the number of parameters that need to be estimated to train one single deep learning model is massively large, so this could also limit the practical use of deep learning. Moreover, to have an accuracy index about stock price or stock return prediction is not enough because the practical utility of the forecast, on the stock market is given by the value of the expected rate of return or expected information ratio. Finally, as previously discussed, deep learning models have a high capacity to overfit the training data, resulting in the low accuracy in predicting the stock price in the test dataset. Our results for the deep learning model show that some machine learning algorithms exhibit the property of overfitting, which is a barrier for stock prediction. It could be thought that Global minimum overfitting is an ill-possed problem of training deep networks. Due to that our results could demonstrate that the predicted stock values and returns are unrealistic, wrongful or unactual.

### 6.3. Future Research Directions

An effort to improve the prediction that is made using neural networks and especially deep learning includes the area model. The training model is built from news data and stock price data gathered from many firms. This collection would be beneficial in a way that could be proven even when considered noisy and the other part of the thesis could be to create a method, or to make appropriate changes that will permit to use outside data in a supervised framework. If the second part of the thesis was successful in using the news data effectively, an addition to the research could be to expand the data inputs, using other datasets e.g weather data, company performance data, consumer data, or company announcements. If a dataset could be built containing some faults in the data, this could also enhance it in another way, and allow building neural network training models that could function effectively under noise conditions. Some simple manipulation to the data will be done in Python code before the models are built, e.g removing technical discontinuities, Filling in missing values, normalisation, or in a particularly noisy section of the data, take a moving average.

## 7. Conclusion

In conclusion, deep learning can be used in stock price prediction as it has close ties to the efficient-market hypothesis, which states that stock prices fully reflect all information in the market but that they do not always do so on a dependable schedule. The connection that seems to exist between alpha and the amount of actionable information not included in historical price data is a promising area to expand on. I hope that open-sourcing the dataset and methodology can lead to a new wave of innovations in this field.

### 7.1. Summary of Findings

The deep learning models improve prediction accuracy in many of the situations and data types that were examined as part of this review. This was the result for both short-term and long-term predictions for singular and multiple stock time series. There was no conjecture, and the various machine learning algorithms and activations functions used in the neural networks were tested at length to determine which produced the most accurate regressions. In the end, the findings show that implementing a deep learning model that includes multiple long short-term memory or LSTMs (Hochreiter & Schmidhuber, 1997) might result in the most accurate predictions of stock prices. Examples like this also prove that our original hypothesis was correct, and that the implications of this are that it is probable that in future more precise predictions can be made if databases expand and more advanced machine learning algorithms are deployed.

### 7.2. Implications and Recommendations

Facilitating stock price prediction with deep learning will most likely continue to improve outcomes in finance. Companies developing stock prediction algorithms likely will benefit most from the base result of the capabilities of selected neural network structures. With more time and data for development and further understanding of new models and data geometries, these companies can effectively find connections among common stock price influencing factors including media, company announcements, quarterly reports, and employee happiness.