Deep Learning Models for Stock Prediction

Stefan Ungurean
ENSC 413/813 Final Project, Simon Fraser University
310316291
sungurea@sfu.ca

Abstract

Stock prediction and market analysis has been at the forefront of investing for the past 80 years. A plethora of methods have been developed over the decades in an effort to beat the market and secure immense profits. Machine learning and deep learning models have shown promise in their predictive power for classification and language processing. However, their ability to accurately predict the market is unknown. This paper highlights the construction of four different machine learning networks and validates them against real-world stock data. The RNN and LSTM networks presented provide promising results in their predictive capabilities and are left to be tested with other stocks and in the real market.

Keywords

RNN – Recurrent Neural Network

LSTM – Long Short-Term Memory

NN – Neural Network

1. Introduction

In the Covid-19 era, the stock market saw a mass influx of day traders and long-term investors buy into the market [1]. This flood of new money drove stocks and the overall market to an all-time high, which is still present today. Whether these new individuals win or lose money, they all buy in with hopes of finding a diamond in the rough. But this story is not unique to new investors, but a common mindset even for large fund managers and long-time shareholders. For years, analysts have searched and created many metrics to best predict stock prices over time to maximize profits and beat the market. However, traditional analytical approaches rely heavily on outside factors such as company valuations, budget reports and other important statistics. These methods have never been extremely accurate and have struggled in their efforts to beat the market. As a result, an accurate tool for predicting stocks and market prices is heavily sought after, raising questions about the potential power deep learning networks may have in creating a predictive algorithm.

Machine learning has increased in popularity and industrial application in the past decade [2]. It has paved the way for unbelievable research in fields like medicine, finance and more. Due to its novelty and impressive power for classification tasks and natural language processing. Naturally, this predictive power can be applied to the market. Several papers have already studied differing algorithms and tested their accuracy. Figure 1.1, below shows an example of the accuracies of different machine learning algorithms. Despite seeming rather low, the accuracies of each network show promise in their predictive capabilities indicating that the correct algorithm hasn't been thought of yet. Consequently, this project aims to analyze the potential of certain machine learning algorithms in correctly predicting market trends. It will start by looking at a simple linear regression model increasing in complexity followed by an investigation into recurrent neural networks.

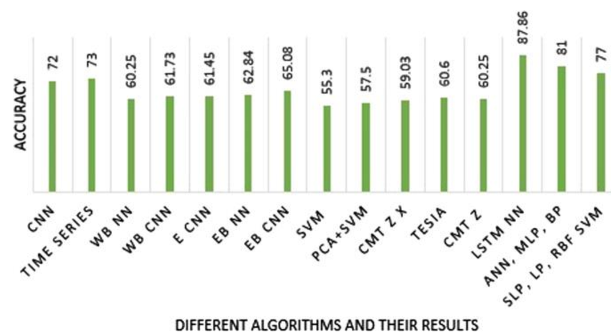


Figure 1.1 Different Algorithm Accuracies on Stock [3]

This paper hopes to use ideas created in the past to develop a new algorithm with accurate stock prediction capabilities.

2. Related Works

There are several papers that have been published in this field. Most notably, is a paper written by Zexin Hu about the use of deep learning models and their prediction accuracy [4]. Here, Hu analyzes models such as CNN's, RNN's and LSTM(Long Short-Term Memory) and compares their advantages and shortcomings. Unfortunately, the paper does not provide a concrete method for highest accuracy.

A second paper worth mentioning by Magnus Olden in 2016 investigates the performance of differing machine learning algorithms on stock prediction. Interestingly, he finds that simple networks outperform deep learning models in accurate stock prediction [5]. This is contrary to the popular belief that more data implies more layers and perceptrons.

Lastly, a paper published by Adil Moghar in 2020 studies the use of LSTM networks in stock prediction [6]. It shows that LSTM networks are powerful in that they use long term data to make predictions. Data from long ago gets lost in regular RNN's, but LSTM fixes these issues and allows that data to have an important say in future predicitons. The paper shows some promising results by using a 4 layered LSTM network followed by a fully connected layer.

Many other papers were published on the subject, however these 3 were the most notable due to their implications in developing the algorithms presented.

3. Data

The datasets used are downloaded from Yahoo finance. They contain the historical stock data with closing prices, opening prices and daily highs and lows. The stocks chosen for the purpose of this paper are GOOGL and APPL. The data contains historical stock data from the past year. This was done to ensure the predictions took relevant data into considerations. Prices from five to ten years ago are insignificant when compared to more recent data. Figure 3.1 shows the historical data of both GOOGL and AAPL stock from July 28th 2020 to July 28th 2021. The models are then tested on data from July 29th 2021 to August 16th 2021. The GOOGL stock was predominantly used to showcase the models in this paper.

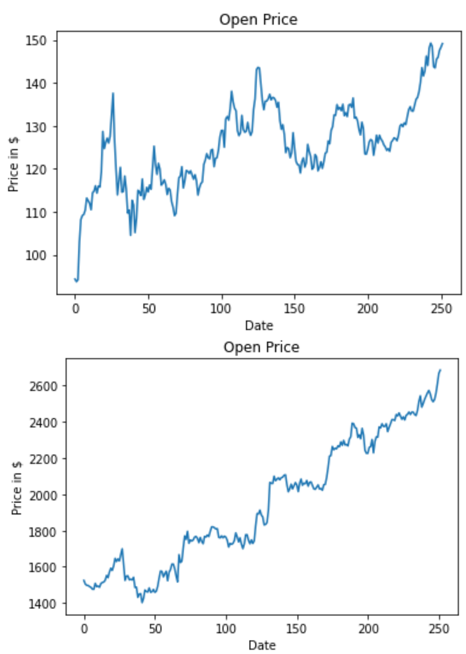


Figure 3.1 Historical Stock Data (AAPL TOP, GOOGL BOTTOM)

3.1. Methodologies

To fully understand which models succeed in stock prediction, four different models were built. First, a simple linear regression with one neuron was implemented. The question here was whether a straight line would be sufficiently accurate in making predictions that could yield positive cash flow from investments. Second was a neural network consisting of multiple fully connected layers. Would a more complex set of linear functions be able to approximate the stock correctly and provide reliable predictions? Third, an RNN was constructed with a fully connected layer and lastly, an LSTM network was built. All models were tested on the test data mentioned above. The statistic for comparison between networks was mean squared error between the actual and predicted values. Since the test data is taken over time, it can be plotted and compared to the model approximations. As a result, MSE is a good indicator of model performance. To build all these networks, the PyTorch python library was used. This allowed for seamless creation and training of all networks and models.

3.2. Regression Networks

The first model consists of a single neuron performing a linear regression on the data. Using an alpha value of 0.2 and training for 2000 epochs, the loss is presented in Figure 3.2 below for the GOOGL stock data.

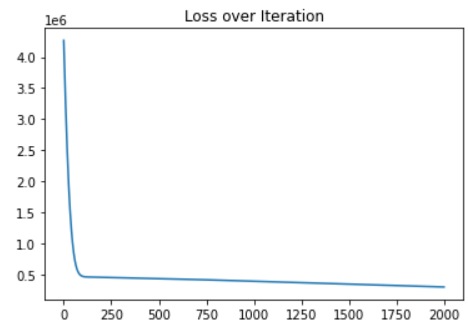


Figure 3.2 Loss Over Iteration, Linear Regression, GOOGL

Figure 3.3, shows the linear approximation.

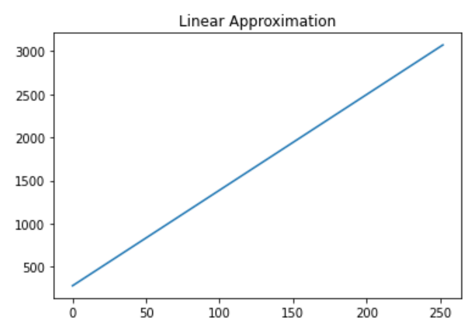


Figure 3.3 Linear Approximation from Regression, GOOGL

Testing this model on the GOOGL test data from July 29th to Aug 16th we see the difference in Figure 3.4.

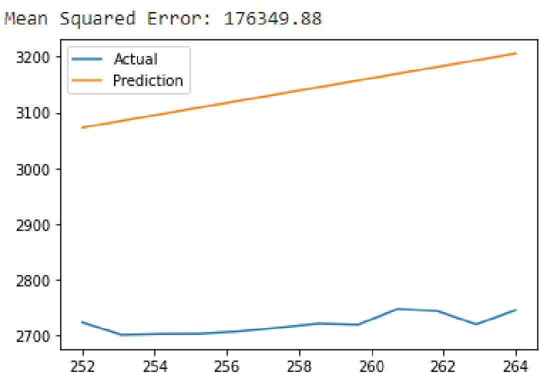


Figure 3.4 Linear Regression Prediction vs Actual Data GOOGL

As shown above, the linear regression model does an awful job of correctly predicting data. Despite the increaser in stock data seeming rather linear in Figure 3.1, the increase is not constant over the course of a few days as our test data shows. The numbers on the x-axis represent the next subsequent 12 days after the training of the model was complete.

The second linear model was slightly more complex. It consists of two fully connected layers of 200 and 100 hidden neurons followed by ReLu activation function. Using a alpha of 0.01 and training for 10000 epochs the loss and overall regression plot is shown in Figure 3.5.

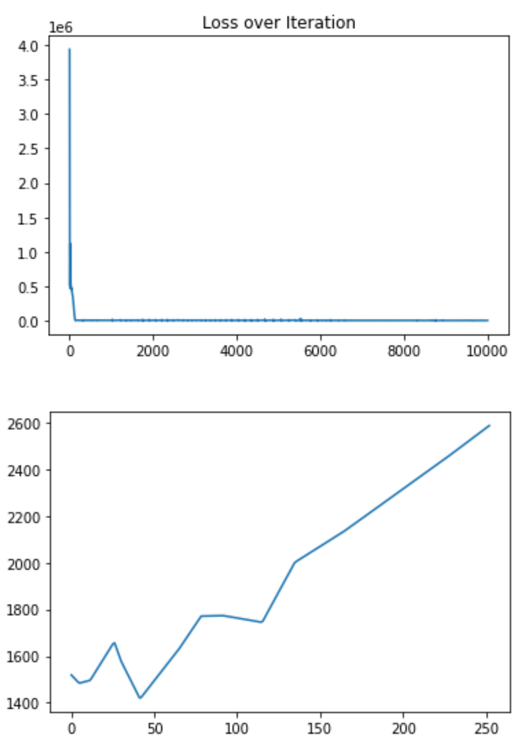


Figure 3.5 Loss (Top), Approximation (Bottom) of Linear Neural Net, GOOGL

As seen by the approximation, it is slightly more complex than its simple linear regression counterpart. It resembles the traditional GOOGL stock more than the previous model. Figure 3.6, shows its performance on the test set.

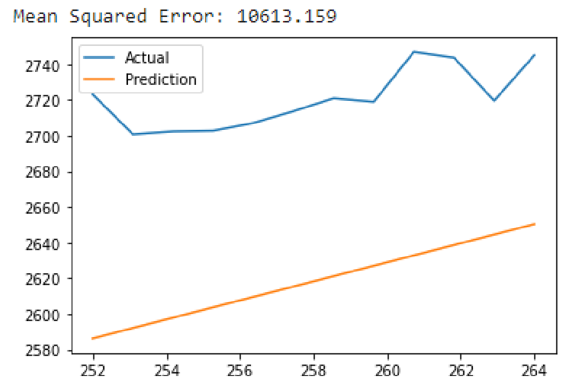


Figure 3.5 Linear NN Prediction vs Actual Stock

As shown above, the MSE is much lower for the neural net than for the simple linear regression, however it heavily undershoots the actual value of the stock by around $100. This is not a great predictive tool.

3.3. RNN's & LTSM

RNN's are powerful tools that allow networks to remember previous data to make predictions for the future. This is crucial for tasks such as language prediction and can be applied to stocks as well. The architecture of an RNN is shown in Figure 3.7 below.

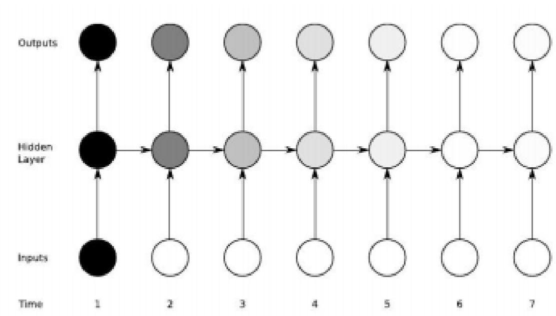


Figure 3.7 Classic RNN Network Representation

Each hidden node takes data from the inputs and the previous hidden node in time. It is a revolutionary idea in training networks to remember data and apply it to sequential problems. However, it runs into an issue. As time progresses, it remembers newer data better than older data and suffers from the vanishing gradient problem. As a result, it tends to make predictions based off the last data in the sequence and does not use earlier sequence data as much.

In order to create proper sequence data for stock prices, the moving window method must be implemented. The moving window takes a certain number of data inputs and uses them as a sequence to predict the next element after the window. This is shown in Figure 3.8 below.

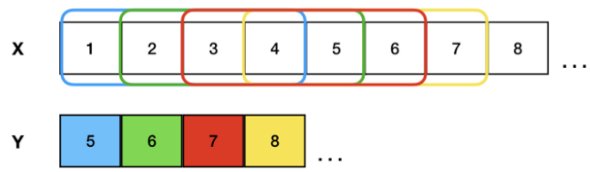


Figure 3.8 Moving Window Method

This creates a sequence of data that can be used as input to the RNN. In this case, the sequence consists of prices over the past 10 days. The RNN takes those sets of sequences as input and compares them to the "11st" day as ground truth. From there it iterates updating the weights as a regular neural network.

The RNN model developed was based off the one created by Moghar and Hamiche [6]. The network consisted of 1 hidden RNN layer with 10 nodes followed by a ReLu activation and a fully connected layer. It also took inspiration from Magnus Olden's paper, simplifying the network as much as possible. Figure 3.8 below shows the loss throughout the epochs.

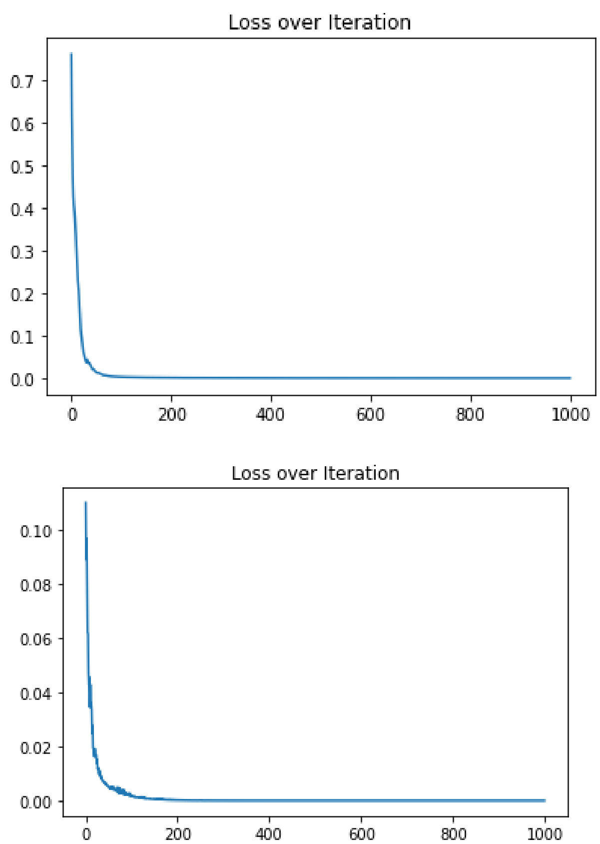


Figure 3.8 RNN Loss Over Iteration Train Data(Top) and Test Data (Bottom)

The model converges very quickly, and the resulting approximation is shown in Figure 3.9.

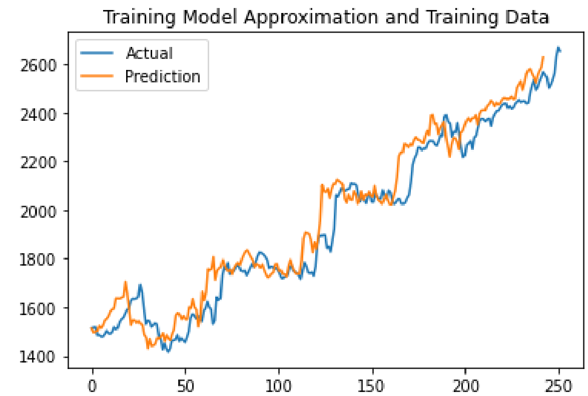


Figure 3.9 RNN Approximation and GOOGL Training Data

As seen above, the approximation is trained quite well with a delay of around 10 days. This is because the data was trained with a sequence of 10 days in mind. Shortening the lookback will improve the approximation but it will not necessarily improve testing accuracy. On the test data the following was plotted.

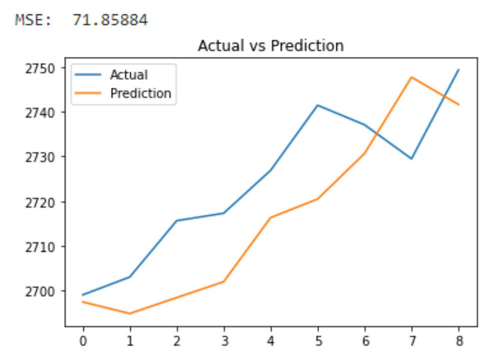


Figure 3.10 RNN Test and Prediction

Figure 3.10 shows an extremely small MSE loss, however the price predictions slightly overshoot majority of the time. This is quite a good model based on the testing dataset.

LSTM

LSTM network are a branch of RNN's that make remembering long-term data easier for neural networks. They are more complicated blocks of operations that allow the network keep track of data that occurred long time ago.

Similar to the RNN network, the stock data had to be converted into a sequence so the same moving window method was used. The model itself consisted of a 2 layered LSTM with 10 hidden nodes, followed by a ReLu activation and a fully connected layer. This was done in a way to mimic Moghar and Hamiche's model as well as attempting to keep it simple via the conclusions from Olden's paper.

The loss of the LSTM network is shown in Figure 3.11 below.

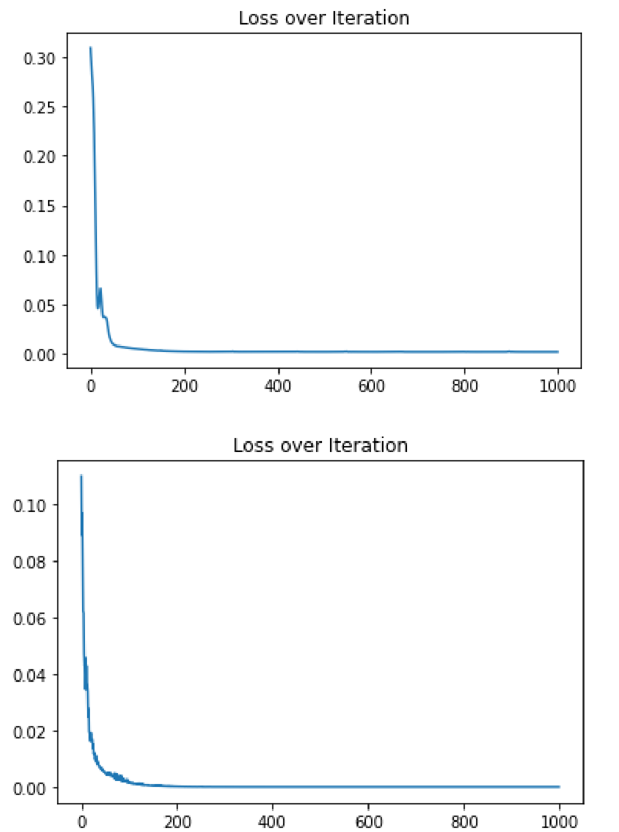
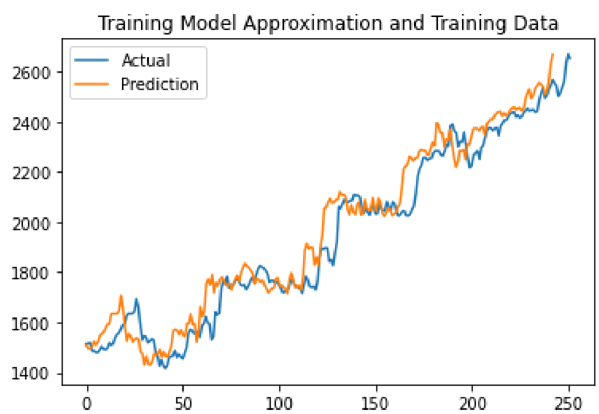


Figure 3.11 LSTM Loss for Training Data (Top) and Test Data (Bottom), GOOGL

Figure 3.12 shows the model approximation and test plot with MSE loss.



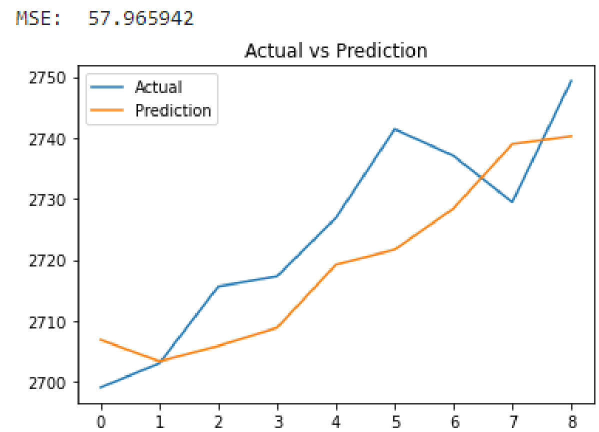


Figure 3.12 LSTM Approximation and Test, GOOGL

Like the RNN network before, the MSE loss is quite low, and the predictions generally undershoot the actual value of the stock. So, by using the predictions, you can make money in the market.

4. Discussion

The four models constructed for this paper were progressively better in correctly predicting stock outcomes. The linear regression and layered linear model were not good predictors of stock data in any way. This was expected as stocks do not follow linear trends on a day-today basis. It was interesting to see that the layered neural network had a lower MSE loss as compared to the linear regression. The network learned certain trends and adjusted the slope of the resulting line accordingly, but the generalization ability was severely lacking. This is due partly since the model suffered from high bias because of few training samples. As a result, both the regression and the neural network did not have good predictive power.

The RNN and LSTM models however, were light years better in making predictions than their linear counterparts. Both models had a very low MSE loss with the LSTM network being slightly lower. Both networks tended to undershoot the actual stock value, and both did not show signs of overfitting via the loss functions from the training and testing dataset. This indicated that the model did not suffer from high bias or high variance. It was important to try and keep the models as simple as possible as per Magnus Oden's findings in his paper while at the same ensuring they were complex enough to generalize correctly. The LSTM model seemed to outperform the traditional RNN by a small margin, however, there is a suspicion that the RNN could outperform the LSTM network if the proper parameters are chosen. The RNN relies heavily on recent data in the sequence whereas the LSTM takes older data into account as well. However, when it comes to stock predictions, past data is simply not as important as recent data. Stocks tend to fluctuate between their previous 5–10-day prices. As a result, data from 6 months prior is nearly insignificant. This leads to the assumption that the regular RNN could outperform the LSTM network.

5. Conclusion

Stock prediction and market analysis has been at the forefront of investing for the past 80 years. A plethora of methods have been developed over the decades in an effort to beat the market and secure immense profits. However, these methods are weak at best since there is no way to accurately predict the rise and fall of prices using traditional market analysis. Machine learning and deep learning models show promise in their predictive power in many areas such as image classification and language processing. By induction, they would be powerful tools in stock market forecasting.

This paper looked at the predictive capabilities of four distinct machine learning networks: a linear regression, a deep linear model, a classic RNN and a LSTM network. The linear models constructed could not generalize accurately due to having high bias and the nature of linear regression models. Day to day stock prices simply do not follow linear trends. As a result, the generalizability of these two models is extremely weak.

The RNN and LSTM network however were much more successful in providing accurate predictions. The LSTM network slightly outperformed the traditional RNN model, but both provided reasonable predictions. Both models tended to undershoot the stock; which is acceptable since it allows for profits in the market. However, further investigation could be done into finding optimal sequence lengths and number of layers to really create precise models.

Future works could perform extensive investigation into optimal parameters for RNN networks that could increase generalizability of the model. Multiple stocks could be considered to create an all-encompassing accuracy score of certain models in their predictive power. Furthermore, actual trading in the market using these models could be done to see if profits would outweigh losses. Hopefully, there are new models discovered that can accurately predict stock prices with a high degree of accuracy.

As of right now, the RNN models constructed for this paper perform extremely well and are left to be tested in a real-world environment.

References

[1] S. Tompor, "Why New Investors Bought Stock During Covid-19," Detroit Free Press, 5 February 2021. [Online]. Available: https://www.freep.com/story/money/personal-finance/susan-tompor/2021/02/05/how-invest-stock-market/4360276001/. [Accessed 14 July 2021].

[2] P. A. Vecina, "A machine learning approach to stock," Barcelona School of Informatics, Barcelona, 2019.

[3] A. Saini and A. Sharma, "Predicting the Unpredictable: An Application of Machine Learning Algorithms in Indian Stock Market.," *Ann. Data. Sci.*, 2019.

[4] Z. Hu, Y. Zhao and M. Khushi, "A Survey of Forex and Stock Price Prediction Using," *Applied System Innovation*, vol. 4, no. 9, 2021.

[5] M. Olden, "Predicting Sotcks with Machine Learning," University of Oslo, Oslo, 2016.

[6] A. Moghar and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," *Elsevier*, vol. 170, pp. 1168-1173, 2020.