Advances in Economics, Business and Management Research, volume 215
**Proceedings of the 2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)**

# Stock Price Prediction Based on Machine Learning: A Review

Kwun Fung Ng[1, *]

[1] *Goizueta Business School, Emory University*
[*]*Corresponding author. Email: alex.ng@emory.edu*

**ABSTRACT**
Designing the optimal machine learning architecture has been an active area of research. A common application of this tool is on the stock price prediction. Putting this in practice raises concern over many aspects—effectiveness, accuracy, and precision. Even if researchers conclude that there is value to attract from machine learning, the question regarding which algorithm to adopt remains. While existing research is dedicated to investigating the accuracy of machine learning, further research sheds light on the advantages and limitations of each model. This article summarizes the classification of machine learning and evaluates the methodology and result of relevant research on applying it to stock prediction under each category. This article also explores some areas for future investigation that tackle crucial shortcomings that would undermine the reliability of the models. The purpose of this work is to offer insights into improving the application of machine learning through various methods of research as well as addressing what has been identified as problems that are common to all algorithms.

***Keywords:*** *Machine Learning, Stock Price Prediction, Regression, Binary Classification.*

## 1. INTRODUCTION

Machine learning has been the primary focus in many industries as the next potential breakthrough. It is applied in practice in many data-heavy jobs and seems to replace human labor by being more efficient and accurate. This idea has also been proven to be pragmatic in preventing fraud, facial recognition, and spam detection. However, the use of machine learning on the stock market is less clear. Trading stocks require a high level of precision and accuracy to ensure long-term net profits. Any miscalculation or misinterpretation of any parameters depending on the model would result in a loss. With a wide range of choices, it is necessary to understand the types of machine learnings (as they are applied differently) and their advantages by investigating the existing research that experiments on the accuracy of each. In this paper, there will be a discussion of the classification system used to categorize machine learning, followed by an introduction to each type.

Related research has been able to identify certain trends among experimental research. In a review written by Strader et al, it was found that multiple parameters lead to better results compared with only one parameter [20]. The same article suggests that abandoning prior understanding of parameters is useful to predict stock prices at high frequency. In the work of Kamley et al, it was suggested that the same algorithm can produce a distinct level of accuracy by integrating various techniques [21]. Some other studies compare the results of these algorithms directly. Isah and Zulkernine states that support vector machine achieves a higher accuracy than K-means after their review in 2019 [22]. All of review summarize experiments that focus on certain aspect of machine learning in stock market prediction. Yet, it should be noted that many of these conclusions do not always yield the same results, and the configuration of their experiments is not specified in the review. To rigorously review research with stock price prediction using machine learning, one must manipulate variables and look at the similarities and differences in their approach.

This article intends to circumvent what related reviews have done and focus instead on the methodology of research across machine learning algorithms. While the work has been done either in general for most algorithms used for predicting stock price, it is unclear how algorithms or a specific one would benefit from new innovations. For example, when investigating the input variables, some authors indicate that multiple parameters are beneficial in predicting stock prices. Yet, the issue in this research method ought to continue and discuss more about whether the trend stays the same with increasing number of parameters in different models. The goal is to provide possible modifications to their research method, rather than a summary of their research findings as existing reviews have been doing. In addition, this work also aims to evaluate the method of comparing results from simulations from which the opposite conclusion may be drawn when the difference is not statistically or consistently significant.

The rest of this paper is organized as follows. In Section 2, within each learning method, some examples of algorithms will be listed, and their corresponding research will be reviewed and examined. This evaluation will focus on the merits of novel ideas and possible improvements in their methodology. The results of the research will also be noted if they provide valuable insights into comparing algorithms. In Section 3, a general review of the shortcomings of machine learning would ideally echo the problems mentioned in the discussion of experiments. The purpose of this section is to address a possible direction for future research at which we hopefully would improve on the issues found in the present understanding of machine learning algorithms. Section 4 concludes the paper.

## 2. DETAILED REVIEWS

Machine learning, within the scope of financial investment or outside this spectrum, does not base itself on one method of learning. This is not a reference to their respective algorithms, but a first step to understanding the general idea of why machine learning could improve over time as it receives a large amount of data. This is crucial since the most suitable type of algorithms would vary depending on what the goal is. While there is no unified theory on how to classify machine learning, it can be put in the following categories—supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

### *2.1. A Review of Stock Price Prediction Based on Supervised Learning*

Supervised learning takes advantages of labeled data—data which is known to represent certain meaning and therefore could be considered as a set of values in a variable. Data are already grouped under supervised learning algorithms. In this case, an input is defined, and a clear output is as well. The process often begins with a large dataset, and a designated function (let it be linear, sigmoid, polynomial, or otherwise). The algorithms take the input, learn its correlation with output, and make predictions on the training dataset [1].

Such mechanism has its value in solving classification and regression problems [2]. The former identifies input values and attempts to put them in categories—that is, “increasing stock price” and “decreasing stock price”, or “overrated” and “underrated”. The latter, on the other hand, creates a model that would be able to use all the inputs and map out the outputs, which ultimately predict the relationship between the two. Often, an accuracy function would determine if the learning is completed (i,e a standard error of 5 unit) [1].

#### *2.1.1. Stock Price Prediction Based on Long Short Term Memory (LSTM)*

Long Short Term Memory, also referred as LSTM, is an algorithm under the supervised learning category. It was developed by Juergen Schmidhuber, a leading expert in the field of artificial intelligence and deep learning [4]. LSTM developed from the Recurrent Neural Network (RNN), which is an alternative to traditional Feedforward Neural Network. LSTM aims to address some limitations of RNN through a better architecture. The two shares the same core principle. The improvement of LSTM that differentiates them is that it has a fast and slow changing output. The latter is designed to deal with certain issues in RNN which would be mentioned in latter sections of this article. In short, it seeks to perform better than traditional RNN when the series is long.

When we focus on existing work that tests the application of LSTM, we see that it performs relatively well. In an experiment conducted by Hengjian Jia, it tests the accuracy of LSTM in stock price prediction by adding a different number of hidden layers and layer sizes [6]. In this experiment, an increase in accuracy is not always consistent with an increase in hidden layer. The authors fail to explain this observation. This offers a new direction for further research. It does not intuitive make sense that increasing hidden layers reduces accuracy. Even if this is the case, this phenomenon should have applied for all layer size. It could be that bigger layer size is prone to vanishing gradients. Another alternative account is that the data must pass through the algorithm multiple times to prevent overfitting. In another study conducted by Sreelekshmy Selvin et al, the team based its model on fundamental analysis, which estimates the share value of a given company by analyzing a range of economic factors such as sales and profits [7]. It also looks at historical stock price average as well as linear and non-linear regression models.

The difference between the results we can obtain from the previous study and this one is that Selvin et al were able to use three machine learning algorithms and compare them. It shows that accuracy is most preferred for CNN (which will be discussed in later sections), followed by RNN and lastly LSTM. The reason, according to Selvin et al, for more accurate results is that it does not depend on previous information, whereas as I explained regarding the mechanism of typical RNN and LSTM, the learning of the other two algorithms would be affected by an output of the previous layer which is passed to the next one. This may be a piece of evidence which shows that CNN is more effective in machine learning. Nevertheless, it could be deduced that CNN is only advantageous in short term prediction when we consider the assumption of technical analysis. This type of analysis assumes that stock prices will behave in an established pattern and repeat itself. This seems to be in line with the other two algorithms (especially LSTM which was already designed to perform better for long series) that take into account the historical information. Since the stock market is highly dynamic, short-term price estimation is rendered more difficult if learning is dependent on previous price points [7]. Nonetheless, all of the results for all three companies and algorithms outperformed ARIMA, which is a linear model for forecasting, implying that the algorithms do have practical values in price prediction.

### *2.1.2. Stock Price Prediction Based on Convolution Neural Network (CNN)*

Convolution Neural Network is perhaps a unique idea compared with the principles of LSTM and SVM. The exact time of the invention has some debate, since it has many precursors or earlier versions. Yet, in 1994, the first CNN occurred, and it was named LeNet5 by Yann LeCun [12]. An appropriate example to illustrate the application of CNN is image analysis and facial recognition. The way to apply this on stock prediction is to analyze stock trends, which can be graphical and therefore passed through convolution layers. Each layer can capture certain features of the image and discover trends. This is particularly useful in determining the rise and fall of stock prices.

In a project called "Stock Prediction Using Convolutional Neural Network", Sheng Chen and Hongxiang He tested the accuracy of applying CNN on stock market prediction [14]. The study used information such as closed opening price, closing price, turnover, and volume of stocks in China. The accuracy of CNN was promising. Iteration is the number of times the image is passed through the algorithm, so this increased the accuracy of the prediction, but the increment was minimal. The conclusion of this project was that CNN was able to predict short term price change accurately. However, some issues could occur such as vanishing gradient, which would prevent the model from learning [14].

The study applied CNN only in binary classification, rather than regression, as noted in the method section of the article [14]. The benefit is that it is easy to achieve a higher than 50 percent accuracy, but it fails to provide actual context regarding whether this is better in terms of predicting prices. Decisions made in the stock market are more complex than merely "buy the stock before it increases in value" and "sell the stock before it loses its value". Many reports and experiments testing the value of machine learning in the financial market have already shown that they are more accurate than simply a random guess. However, to actually have broad use of the algorithm, it must show precision—the degree to which it is able to map the trends of stock prices. A better methodology should surround regression, rather than classification, and examine the algorithms that perform best in terms of the difference between the predicted and the actual. Sidra Mehtab and Jaydip Sen in 2020 focused on stock price prediction using CNN and LSTM-based deep learning models. It used two algorithms to predict the same set of data. The data includes two weeks worth of data from NIFTY 50, which is a benchmark representing the weighted average of India's largest companies on the National Stock Exchange.

Mehtab and Sen concluded that both algorithms showed high level of accuracy and precision [15]. Nevertheless, the LSTM model outperforms CNN by a slight advantage. LSTM also responds quicker than CNN, which is crucial since the conditions and actions in the stock market change by seconds. While the study does imply that LSTM performs better in stock market predictions, it is not clear if this is statistically significant. The difference between the two algorithms' precision is not too far from each other. To conclude with greater confidence, it would be ideal to run the model with more than one configuration as the number of iterations can cause a difference in accuracy. Therefore, this parameter may cause the result to be the exact opposite. In other words, this is not a conclusive study to say that CNN is an inferior predictive tool for stock prices. Rather a future study is needed to improve this hypothesis with regards to iterations.

## *2.2. A Review of Stock Price Prediction Based on Unsupervised Learning*

Unsupervised learning is by definition the opposite of supervised learning. The dataset is not labeled. There is clear variable to predict. If one is to consider supervised learning is a process to find Y=f(x), then unsupervised learning is solely exploring x [3]. There is no correct answer as we did for supervised dataset where we know the true answers for each set of input values. Now, there is no training dataset that teaches the algorithms to find what we seek. Instead, a relationship between inputs is investigated [1]. Problems associated with clustering and association tend to find this tool useful. Since the values in a dataset are not labeled, the algorithm would attempt to find hidden groups within the dataset. An example would be finding customer groups by investigating their purchasing record [3]. An association problem deals with understanding commonalities between inputs. An example to illustrate this is that if one is taller than average, then one is also heavier than average.

Researchers would use this principle to find common properties that lie in the dataset.

### *2.2.1. Stock Price Prediction Based on K-Means/Hierarchical Agglomerative Clustering (HAC)*

K-means is a relatively common algorithm used to tackle classification issues. Its principle surrounds the idea resembling support vector machine which is another algorithm to be discussed later. It seeks to define clusters of data points which are categories in a large dataset. Depending on the dataset itself, the categories may be customer types or stock types. The algorithm would randomly choose the coordinates for a centroid which is the center of the cluster, compute the distance between the data points and the centroid, and group them based on minimum distance to the centroid [19]. It does this repeatedly until it finds the optimal groupings given a certain number of centroids determined before the algorithm begins. HAC is an algorithm that shows how the individual data points are connected. It assumes that all data point are a singleton. The first step is to find the two singletons of the shortest distance. Then, the two points for a cluster. Then, the process repeats until there is only one cluster [19].

In Babu, Geethanjali and Satyanarayana's article, researchers investigated K-means and HAC in stock trading. It sets the number of clusters as 2, each representing rise and fall. Then, if the algorithm predicts a rise, a decision of buying the stock is determined. It combines the two called Hierarchical agglomerative and Recursive K-means Clustering (HRK). The result was that the accuracy of predicting the rise and fall of stock is consistent across K-means, HAC, and SVM (to be discussed in the next chapter), but HRK appears to outperform K-means in creating profits. The article includes both numerical information from financial statements of companies and qualitative information from the news that are categorized into 0 and 1. The authors did a controlled experiment that quantitatively compares results across HRK. HRK which takes both qualitative and quantitative measure was able to obtain the most profits. I argue that existing machine learning algorithm should focus on quantifying qualitative measures. The reason lies in the fact that the stock market is full of investors who are affected by expert predictions, news release regarding corporate decisions, and a general sense of company's image. By taking these into accounts such as improvement, complaint, and reorganize, the algorithm would best mimic trading in the stock market. Machine learning in the financial industry, then, is no longer a product and tool of financial mathematics; it is a beneficiary of the social sciences.

## *2.3. A Review of Stock Price Prediction Based on Semi-Supervised Learning*

This particular category lies between the previously discussed two. A large amount of input values is present, only some of which is labeled. Algorithms in this category serve to organize data and at the same time make predictions. They learn from a training dataset and aim to correctly predict the output of the given inputs [1]. The situations in which this becomes advantageous are more commonly seen in real-life where some data are labeled, and some aren't. Photos, for instance, are often partly labeled. They may be put in categories such as animal, human, and buildings. However, some might come with only the photo itself. Then, the goal is to learn from what has been labeled and classify the unlabeled by learning from the training dataset [3].

### *2.3.1. Stock Price Prediction Based on Support Vector Machine (SVM)*

Support Vector Machine, or SVM for short, is within the category of semi-supervised learning [20]. It was introduced by Vladimir Vapnik in 1989 [8]. SVM has been improved and modified many times in unique ways after its birth. Unlike the principles of RNN or LSTM, it does not focus on generating a particular output, but instead it seeks to produce a line between data of two categories. Therefore, SVM is mostly used for solving classification problems instead of regression. The goal of SVM is to find the optimal line in a hyperplane—optimal in the sense that it maximizes the margin of the line separating the data points from two classes [9]. When it is applied to the stock market, the two classes would become rise and fall of stock prices.

Research carried out by Fenghua Wen et al saw the potential with SVM on the financial market and therefore tested the accuracy of SVM in 2014 [10]. It not only tried to apply a traditional algorithm of SVM but incorporated method of analysis to better turn historical stock price into an identifiable trend. The conclusion of the research lauds the accuracy of the algorithms [10]. The significance is that with better analysis that can be integrated into the model, the prediction would yield a value closer to the actual one. Thereupon, there are differences between the predicted values, even when the same algorithm is used, indicating signs of improvement with better theoretical understanding of stocks.

Outside the paper's conclusion, we may begin to see that the stock market is best predicted with non-linear models. For regression-based algorithms, seldom is the chosen or resultant model linear or polynomial. It is usually a more complex one. Singular Spectral Decomposition (SVD) is an improvement upon traditional SVM. This method is incorporated to analyze non-linear and non-stationary time series [10]. This allows SVM to better model the behavior of stock price fluctuations which then turns into better prediction results. This is intuitive correct as we seldom see a very identifiable pattern in the diagrams of any stocks. When we continue to use machine learning algorithms to predict prices, we should incorporate techniques such as this to avoid the assumption that the market is linear. Even for identifying customer behavior, customers do not belong to the same demographics.

In another research conducted by Junyoung Heo and Jin Yong Yang, a number of values listed on the company's financial statement including earnings per share and net profit of 200 companies listed on KOSPI 200 was analyzed [11]. The study aimed to provide a comparison between expert prediction and SVM prediction with various combinations of the values the study included. The conclusion Heo and Yang gave was not about whether SVM itself is accurate but rather that machine learning under SVM can sometimes be more effective than human analysis [11]. In addition, more parameters do not always guarantee better results than a single parameter. This suggests that Strater et al's conclusion is not true for any number of parameters.

In this particular report, SVM is able to predict around 55% of stock price fluctuations. Comparing the results from Babu et al which has a similar level of average accuracy for the SVM samples, we can see that HRK is a better algorithm. This was attributed to HRK's inclusion of qualitative features, the combination of two clustering methods, and a proper number of splits of clusters. Perchance, other than including the appropriate parameters for learning, improvements to be made on existing algorithms in the stock market should attempt to combine algorithms. A possible idea would be running a CNN on a given stock's price trend to examine its rise and fall, followed by a comparison with the result of non-linear SVM using SSA technical applied in the work of Fenghua Wen et al.

### *2.4. A Review of Stock Price Prediction Based on Reinforcement Learning*

Reinforcement Learning is based on the idea of trial and error. Similar to supervised learning, feedback is provided. The difference is that the feedback is not right or wrong but rewards and punishments for a set of actions. It seeks to maximize rewards by performing the right task and change the behavior when the action is punished [1]. This learning mechanism is usually applied in robot learning. Instructions or a goal is given, and the robot would attempt to reach that goal by performing a set of actions which are constantly corrected until it finds the optimal behavior [1].

Present reinforcement learning strategies mainly depend on outputs of supervised, semi-supervised, or unsupervised learning algorithms. A reinforcement algorithm such as Q-learning is placed in the next layer to self-direct available options and optimize its decisions. To improve on existing application of machine learning requires more efficient use of the algorithms discussed in earlier chapters. However, by adopting a second layer of reinforcement learning, it may yield interesting results. The degree of rise and fall may allow the model to explore arbitrage opportunities.

## 3. SHORTCOMINGS OF MACHINE LEARNING

### *3.1. Overfitting/Underfitting*

Overfitting is a shared issue in many of the existing algorithms. This concept posits that the algorithm would create a model that overly adheres to the training dataset. While the result may appear that the model can accurately predict the results of the training dataset, it does not produce the optimal model [16]. There will be outliers, also known as noises, in a training dataset. When the algorithm includes these noises, the model becomes distorted and does not apply the most accurate function to predict new data. Many current work focuses on creating an algorithm that could mitigate this issue, because this issue is a direct influence on the accuracy of the resulting model. Highly flexible algorithms tend to be most vulnerable to overfitting. It may create a model that attempts to pinpoint all data points of the given dataset, but this also implies that such learning would not be able to generalize its results to anything outside the training dataset [16]. Less sophisticated algorithms, on the other hand, are more resilient to overfitting. Because of their structure, they remap the training dataset with a lower accuracy. There is a trade-off between accuracy and the degree of overfitting that results in the issue of generalization. Much effort is put in to find this balance.

To exemplify, in SVM's instance, a number of points colored in blue or purple is spread on a two-dimensional space. To classify the points, a perfect curled line may be drawn to separate them. An alternative to that would be a straight line which generalizes the pattern. While the former may produce the best result for the dataset, only the latter would ensure minimal error in the long run as not all datasets would behave as the training dataset. SVM has a fast and slow changing component in each layer that both learns and remain distant from being too affected by each data point.

Underfitting is a less significant issue in machine learning. It means that neither the training dataset nor new data is modeled, which contrasts to the issue discussed above [16]. This is often a consequence of choosing the incorrect model or algorithm. This is resolved through seeking and adopting alternatives.

### *3.2. Lack of Representative Data*

Most algorithms, supervised or otherwise, need some sort of data to train the model. The algorithm could only do so much as to learn from what it has. When the data is noisy, not representative, or incomplete, the algorithms would produce a less desirable model [17]. Noisy data tend to move the model away from becoming an optimal predictor. Incomplete data would mean that there is less parameters of an input than needed to train the model. In the case of predicting stock prices, historical data may not be as representative as one may believe as the stock market is highly dynamic, and the structure changes over time.

## 4. CONCLUSION

Most algorithms, supervised or otherwise, need some sort of data to train the model. The algorithm could only do so much as to learn from what it has. When the data is noisy, not representative, or incomplete, the algorithms would produce a less desirable model [17]. Noisy data tend to move the model away from becoming an optimal predictor. Incomplete data would mean that there is less parameters of an input than needed to train the model. In the case of predicting stock prices, historical data may not be as representative as one may believe as the stock market is highly dynamic, and the structure changes over time.

After analyzing the commonalities and differences between existing research on machine learning algorithms, it is not difficult to see that there several directions between experiments investigating on accuracy, precision, comparison and/or improvement. Most research has shown that there are potentials for implementing machine learning algorithms. More valuable conclusions stem from the further research looking at the compatibility of algorithms with stock price prediction. Some algorithms, perhaps due to the fact that they are, on principle, more adherent to the dynamics and understanding of stock market prediction, yield better results in their predictions. However, advantages of a given algorithm are not consistent in all studies. In the work of Sreelekshmy Selvin et al, CNN appears to be more accurate than LSTM. Yet, Mehtab and Sen suggest that LSTM actually outperforms CNN by a slight advantage. The more decisive factor to conclude which algorithm to use depends less on the algorithms themselves and more on their configurations. Moreover, further research should not adopt a binary classification, since this is not a realistic representation of the information needed to beat the market. In fact, other factors neglected such as the transaction cost should be included as a determinant. It would also be an option to construct a full simulation of the stock market with option trading, securities, and selling short. With reinforcement training, machine learning might act as a human and actually be more profitable.

As seen in research conducted by Fenghua Wen et al, improvements could be made by integrating better techniques into the algorithms. This indicates that implementing machine learning into practice is not so simple as many factors are involved. To fully see with context what each algorithm can do at present, there must be holistic experimentations comparing algorithms using the same set of data under the same period, setting the configurations as optimal, and integrate techniques that is appropriate to the respective algorithms. Nevertheless, the primary focus and attention need not be on finding the maximum accuracy and precision. As seen in Heo and Yang's work, machine learning, assuming the algorithms are similar in terms of their performance which is the trend seen in this article, is not significantly better than expert predictions. Therefore, replacing much of stock price prediction with merely algorithms is still too soon. Not until there is an advancement in their effectiveness, will we entirely entrust the work of analyzing stock prices that is traditionally carried out manually on machines.

## REFERENCES

[1] Armstrong, H. (2015, July). *Machines that learn in the wild*. Nesta. Retrieved November 19, 2021, from https://media.nesta.org.uk/documents/machines_that_learn_in_the_wild.pdf.

[2] IBM Cloud Education. (n.d.). *What is supervised learning?* IBM. Retrieved November 19, 2021, from https://www.ibm.com/cloud/learn/supervised-learning#block-ibm-cloud-content.

[3] Brownlee, J. (2020, August 20). *Supervised and unsupervised machine learning algorithms*. Machine Learning Mastery. Retrieved November 19, 2021, from https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/.

[4] Katte, A. (2020, March 14). *Meet Juergen Schmidhuber - the father of LSTM, an outsider in the world of Deep learning*. Analytics India Magazine. Retrieved November 19, 2021, from https://analyticsindiamag.com/meet-juergen-schmidhuber-the-father-of-lstm-an-outsider-in-the-world-of-deep-learning/.

[5] Olah, C. (2015, August 27). *Understanding LSTM networks*. Understanding LSTM Networks -- colah's blog. Retrieved November 19, 2021, from https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[6] Jia, H. (2016). Investigation into the effectiveness of long short term memory networks for stock price prediction. *arXiv preprint arXiv:1603.07893*.

[7] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 international*

*conference on advances in computing, communications and informatics (icacci)* (pp. 1643-1647). IEEE.

[8] Guyon, I. (2016, July). *Data Mining History: The invention of Support Vector Machines*. KDnuggets. Retrieved November 19, 2021, from https://www.kdnuggets.com/2016/07/guyon-data-mining-history-svm-support-vector-machines.html.

[9] Pupale, R. (2019, February 11). *Support vector machines(svm) - an overview*. Medium. Retrieved November 19, 2021, from https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989.

[10] Fenghua, W. E. N., Jihong, X. I. A. O., Zhifang, H. E., & Xu, G. O. N. G. (2014). Stock price prediction based on SSA and SVM. *Procedia Computer Science*, *31*, 625-631.

[11] Heo, J., & Yang, J. Y. (2016). Stock price prediction based on financial statements using SVM. *International Journal of Hybrid Information Technology*, *9*(2), 57-66.

[12] CULURCIELLO , E. U. G. E. N. I. O. (2017, April 18). *The History of Neural Networks*. Dataconomy. Retrieved November 19, 2021, from https://dataconomy.com/2017/04/history-neural-networks/.

[13] IBM Cloud Education. (n.d.). *What are convolutional neural networks?* IBM. Retrieved November 19, 2021, from https://www.ibm.com/cloud/learn/convolutional-neural-networks.

[14] Chen, S., & He, H. (2018, October). Stock prediction using convolutional neural network. In *IOP Conference series: materials science and engineering* (Vol. 435, No. 1, p. 012026). IOP Publishing.

[15] Mehtab, S., & Sen, J. (2020, November). Stock price prediction using CNN and LSTM-based deep learning models. In *2020 International Conference on Decision Aid Sciences and Application (DASA)* (pp. 447-453). IEEE

[16] Brownlee, J. (2019, August 12). *Overfitting and underfitting with machine learning algorithms*. Machine Learning Mastery. Retrieved November 26, 2021, from https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/.

[17] Ohri, A. (2021, March 6). *17 common issues in machine learning: Simplified*. Jigsaw Academy. Retrieved November 26, 2021, from https://www.jigsawacademy.com/blogs/ai-ml/issues-in-machine-learning.

[18] Ding, S., Zhu, Z. & Zhang, X. An overview on semi-supervised support vector machine.*Neural Comput & Applic* **28,** 969–978 (2017). https://doi.org/10.1007/s00521-015-2113-7

[19] Babu, M. S., Geethanjali, N., & Satyanarayana, B. (2012). Clustering approach to stock market prediction. *International Journal of Advanced Networking and Applications*, *3*(4), 1281.

[20] Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y. H. J. (2020). Machine learning stock market prediction studies: Review and research directions. *Journal of International Technology and Information Management*, *28*(4), 63-83.

[21] Kamley, S., Jaloree, S., & Thakur, R. S. (2016). Performance Forecasting of Share Market using Machine Learning Techniques: A Review. *International Journal of Electrical & Computer Engineering (2088-8708)*, *6*(6).

[22] Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies*, *7*(2), 26.