# Stock price prediction based on CNN model for Apple, Google and Amazon

Xiaojian Zhang[1, *]

[1]Department of Economics, University of Birmingham, Birmingham, UK

*Corresponding author: xxz905@student.bham.ac.uk

**Abstract.** The price movements of stocks directly affect the economic interests of investors as well as influence and reflect the macroeconomic policies of the country. This paper initially describes the CNN model's development and fundamental composition before proposing a method for stock prediction based on the CNN model and using it to analyze data from Apple, Google, and Amazon. According to the analysis, it is indicated that all three businesses stocks will decline going forward, with Apple and Google's decline being greater and Amazon's decline less. In the upcoming months, it's anticipated that stock prices will increase a little bit and swing between $100 and $125 a share. The RNN model was implemented to compare the findings in the end, and both models produced about the same stock forecast trend. These results shed light on guiding further exploration of stock price forecasting in terms of the state-of-art machine learning scenarios.

**Keywords:** Stock price prediction; CNN; RNN; machine learning.

## 1. Introduction

Stock forecasting is the process by which a stock analyst with extensive stock knowledge predicts the future direction of a company's stock and the variation degree depending on the development of the stock market. This kind of predicting is based on presumptions that already exist. Making an accurate stock price prediction might result in substantial earnings [1]. According to the efficient market theory, stock prices accurately represent all information that is currently known [2].

The government values the stock market's position as a barometer and an alarm for the economy and the mood of shareholders, as well as for investors. As they say, everything is done in advance, yet nothing is done in advance. For investors in stocks, there is a direct correlation between the ability to predict market trends and the ability to make money. For publicly traded companies, the stock index reflects the operation and future development trend of the company. Besides, it also affects the effectiveness of a company and is the main technical indicator used by companies to study and analyse the market. In other words, the more accurate the forecast, the more certain the prevention of risks. The study of stock forecasting is equally crucial to the nation's economic growth. As a result, the field of research of the stock market's intrinsic value and forecasting has significant theoretical relevance and future practical possibilities. Stock forecasting is a time series and there are many stock forecasting methods to choose from (e.g., LSTM, SVM models), but none of these methods are perfect [3]. For example, SVM models can perform well on both low and high-dimensional data (i.e., a few features and many features). However, when the sample size exceeds a certain number and the kernel function is mapped to high dimensions, the SVM becomes computationally overloaded. In this case, it becomes less suitable for usage and the model is sensitive to missing data, which is often the case with stock forecasts [4]. Although predicting stock prices is indeed a problem that has still not been solved. The fact is quite simple: a lot of factors influence a stock's price, and the historical price of a stock is only a small part of the many reasons. Thus, predicting the movement of share prices is a very difficult problem. In additional, according to Pawar et al., which is an RNN-LSTM is suggested to predict the closing price of the stock the following day. Support vector machines, regression, random forests and neural networks are examples of common machine learning techniques that are compared to the model. They discuss how customers' sentiment affects stocks as the trend changes, and the results have relatively high accuracy, but the model still has some shortcomings. For example, it fails to incorporate sentiment factors such as news and national policies into the prediction [5].

In this paper, the CNN model is chosen to predict the stocks of three tech companies, Apple, Google and Amazon respectively. The reason why this paper is predicting these three companies is because in the age of information technology will probably lead the world in progress and these three companies are leaders in their respective fields and are more representative of the technology stocks. The rest part of the paper is organized as follows. The Sec. 2 will present the origin and basic structure of the CNN model, as well as the visualization of stock data from three companies. The Sec. 3 will predict for these companies will be made using the CNN model and compared with the results of the RNN model, finally limitations to these models are presented. The Sec. 4 will summarize the paper content.

## 2. Methodology

The Convolutional Neural Networks (CNN) model is a special case of DNN, which is often used for image recognition, but is now also used in other fields such as NLP. Firstly, a brief introduction to the DNN model. Fig. 1 gives a typical structure of the DNN model. To be specific, the hidden layer, output layer, and input layer are the three various sorts of layers that make up the deep neural network. The input layer is where the data is sent, and the hidden layer is where the calculation is done with the data. The adjective "deep" denotes the presence of several hidden layers [6].

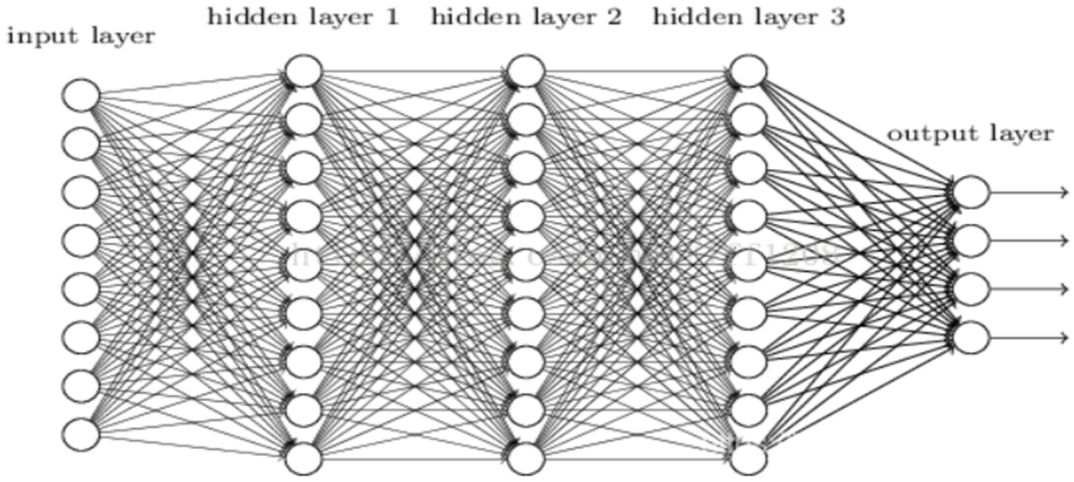


**Fig. 1** DNN model structure.

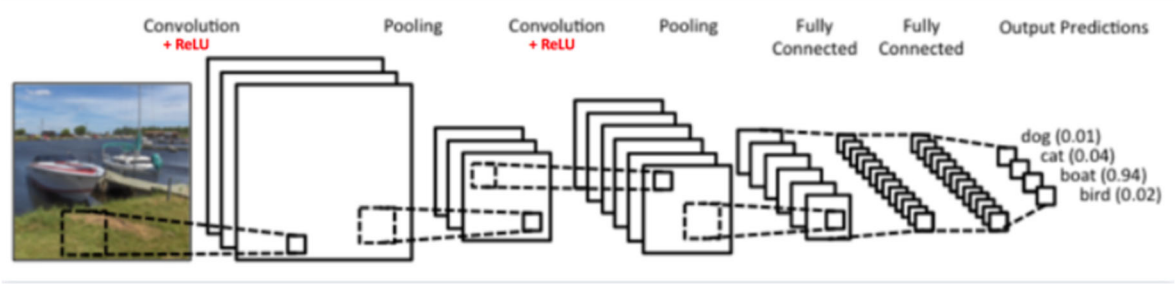


**Fig. 2** CNN model structure.

A CNN model for image recognition is illustrated in Fig. 2. The leftmost boat is our input layer, which the computer can understand as a number of matrices. Next is the Convolution Layer, which is specific to CNN. The activation function in the convolution layer is ReLU and the pooling layer follows the convolution layer and is likewise exclusive to CNNs; it does not serve as an activation layer. As shown in Fig. 2, there are three layers total: two convolutional layers, a pooling layer. Nevertheless, the number of layers might change based on the model's requirements and either a convolutional layer plus another convolutional layer or a convolutional layer plus a pooling layer can be used. A Fully Connected Layer (FC) is followed by a number of convolutional layers plus pooling layers, which is actually a DNN structure. where the convolution of two dimensions in a CNN is defined in the CNN as

$$S(i,j) = (X * W)(i,j) = \Sigma_m \Sigma_n x(i+m, j+n)\omega(m,n) \quad (1)$$

where W is the convolution kernel and X is the input. If the input matrix for X has two dimensions and W is a two-dimensional matrix. W is a multidimensional tensor if X is a tensor with multiple dimensions. X and W maintain their original sizes. For the pooling layer, which is simpler than the convolution layer, pooling is the compression of each submatrix of the input tensor. In the case of 2×2 pooling, then every 2×2 elements of the submatrix is turned into one element, thus reducing the dimensionality of the input matrix.

A pooling criterion is required if one converts each of the input submatrix's n x n elements into a single element. There are two standard pooling requirements. For the pooled element value, MAX and AVERAGE, respectively, utilise the maximum or average value of the associated area [7]. In this forecast, AAPL [8] GOOG [9], AMZN [10] daily historical stock price data from November 1, 2014 to July 10, 2022 were collected as the basis for forecasting through Yahoo Finance. This stock prediction will be trained using the simplest convolutional neural network. For processing time series data with CNN, a one-dimensional convolutional network Conv1d is usually used; the structure of this experiment is: the convolutional layer expands the one-dimensional data into a three-dimensional tensor through a 2×2 convolutional kernel, discards the data less than 0 using the activation function ReLU, then converts the three-dimensional tensor into a one dimensional tensor using a fully connected layer, and finally obtains the final prediction by two Linear transformations.

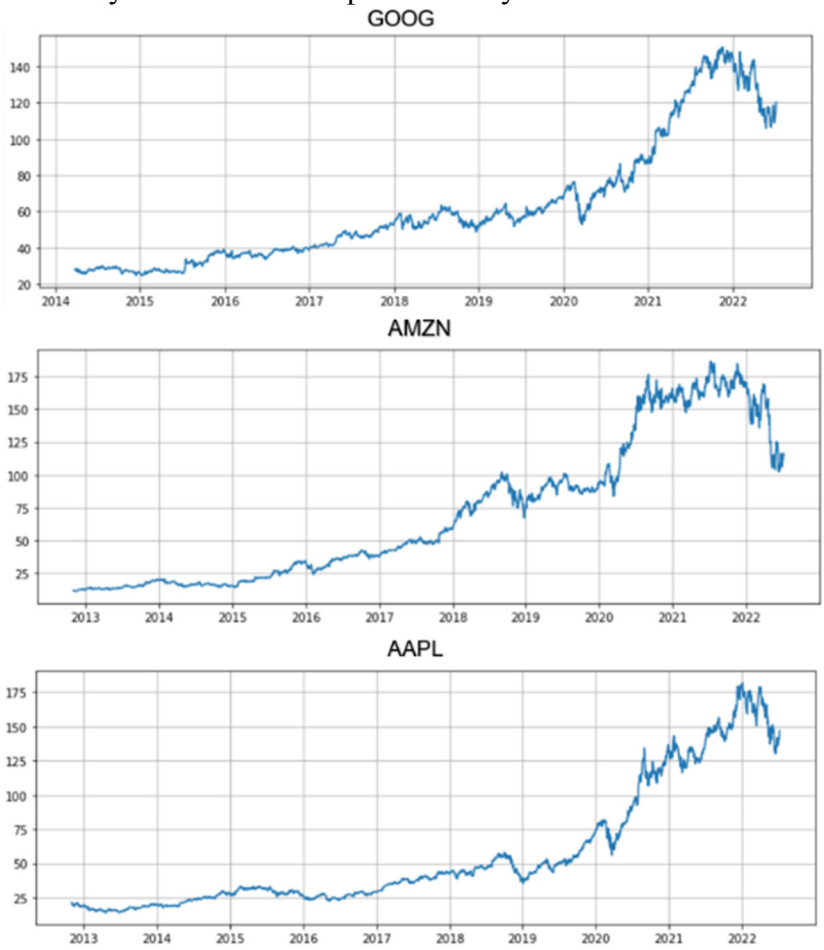


**Fig. 3** The stock price of Google, Amazon and AAPL from the upper panel to the lower panel.

## 3. Results & Discussion

First, one needs to load the historical stock data of the three companies and visualize the data to get the following three charts. The Fig. 3 exhibits the daily closing prices of Apple, Google and Amazon. Accordingly, the stock prices of these three companies have increased year on year and

have increased rapidly up to around $160 per share around 2020, but after 2022, the stocks of all three companies have fallen to varying degrees.

Firstly, the data must be normalised, and then the training set and a validation set must be separated. In this scenario, normalising the data to (-1,1), which is the statistical coordinate distribution, will help generalise the statistical distributivity of the samples. The output of the sample is normalised since the neural network is trained (probability calculation) and forecasted in terms of the statistical relative probabilities of the sample in the event, and the Sigmoid function. accepts values between 0 and 1. A training set and a validation set are created from the data set. A group of samples used to fit parameters is known as a training set. The current model serves as the trainable weight parameter during the training phase, making predictions for each sample in the training set and comparing them to the target. On the other hand, the validation set is a unique group of data reserved for the model training procedure that may be used to tweak the model's parameters and do a brief summary of the model's performance. Typically, the training, validation and test sets are divided into 6:2:2 when the amount of data is not very large (below 10,000) [11]. Taking the apple data as an example, there are 2808 sets of data in this experiment, and the last 170 sets of data are used as the validation set for prediction and comparison with the real values. The outcome of the dataset was split into a training set and a test set.

Based on what is mentioned in the methodology, a CNN model is built and the stock prices of Apple, Google and Amazon are predicted according to the results of running the model, the predicted share price for the last 171 days was compared with the true value. The final predicted stock price for 2022/08/01 is 103.99 for Apple; 92.69 for Google; and 100.69 for Amazon. The following 3 charts are visualisations of the predicted stock prices for each of these three companies

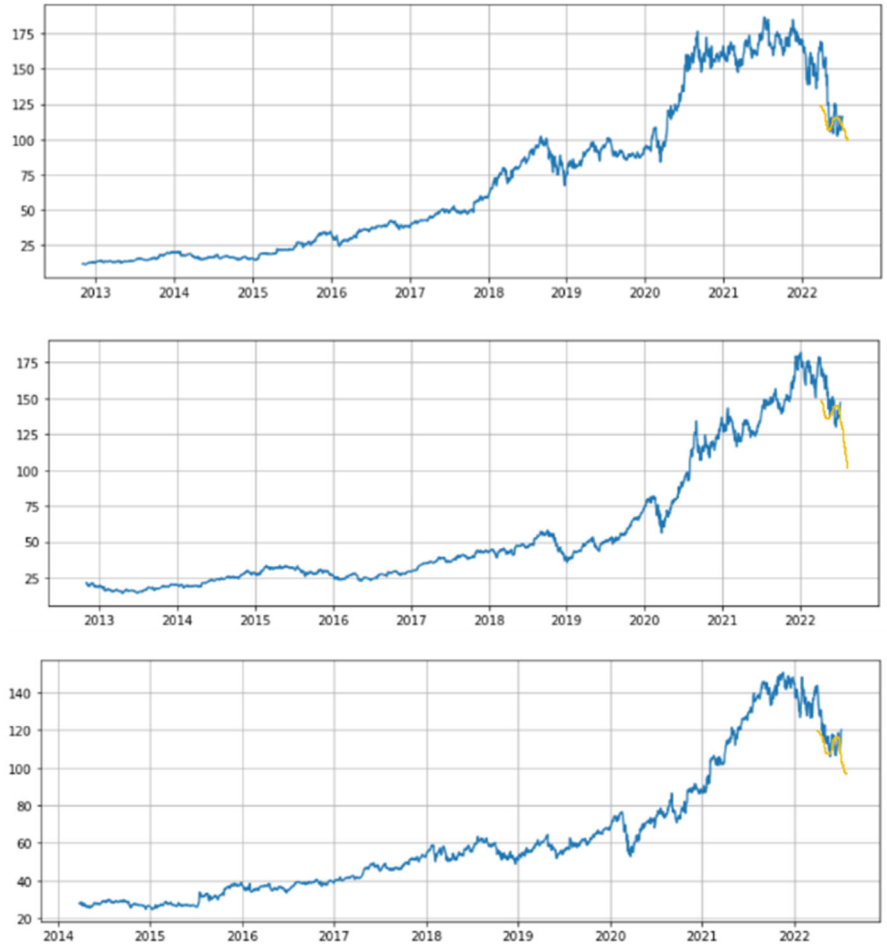


**Fig. 4** The prediction results of Google, Amazon and AAPL from the upper panel to the lower panel.

The charts presented in Fig. 4 show that all three companies' stocks are trending lower compared to historical stock price data. It should be noted that Apple and Google are predicted to drop more,

forecast to fall to approximately $103 per share and $92 per share respectively by August 1, 2022, while Amazon is down less, expected to float between $100-$125 per share thereafter. Overall, though the stock is trending down, the share price will rise slightly in the second half of the year and float around $125 per share.

Subsequently, this paper will compare the results of this model using RNN approach and provide an explanation. The RNN model is another form of DNN in which a neuron's output can denotes that the input of a neuron in layer I at instant m includes both its own output at that time and the output of a neuron in layer (i-1). will also affect itself at the next timestamp. The basic structure diagram is depicted in Fig. 5. One can see the links between nodes in the hidden layer being added. Expanding the RNN model over the time series, one obtains the Fig. 6, where the network's current (t+1) ultimate findings. O(t+1), which represents the outcome of the input and all previous data at that moment., this serves the purpose of modelling the time series. Firstly, this operation will build the neural network using sequntial. The first recurrent layer contains 80 memes, with each time step pushing h$t$ to the next layer, using a dropout of 0.2; the second recurrent layer contains 100 memes, with only the last time step pushing h$t$ to the next layer, using a dropout of 0.2. Since the output value is the opening price on the 61st day. Secondly, the training of the network model will be performed, and the relevant settings in the fit are, in order, training set samples, training set labels, small batch size 64, training rounds 50, test set, training set loop 1 round for one test, and execution of breakpoint training. The training and validation sets are then subjected to loss function operations in this experiment. The model's generalisation performance increased and the prediction accuracy increases with decreasing loss functions. The results are also visualised for Apple, Google and Amazon, respectively.

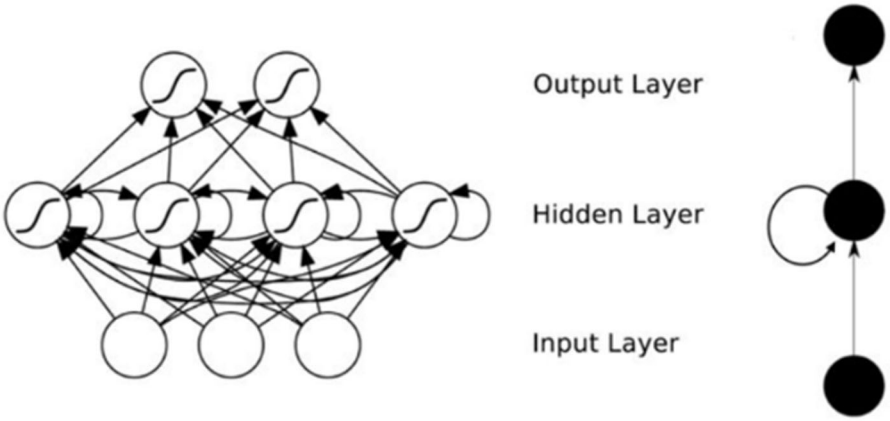


**Fig. 5** RNN model structure.

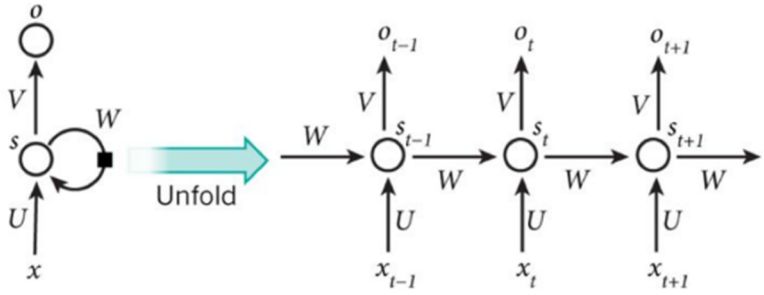


**Fig. 6** RNN model over the time series

As can be seen from the Fig. 7, the error is not significant and the operation can continue. The final prediction is obtained in Fig. 8. After visualizing the results, three metrics are adopted to analyse the prediction results, i.e., MSE (Mean Square Error), RMSE (Root Mean Square Error), and MAE (Mean Absolute Error). According to the analysis, the following conclusions have been obtained, where AAPL has MSE of 19.98, RMSE of 4.47, MAE of 3.75; GOOG has MSE of 10.82, RMSE of 3.29, MAE of 2.71; AMZN has MSE of 15.66, RMSE of 3.96, MAE of 3.038.

According to the results, it can be seen that when using the RNN model, there is an error, but its value is not very large, i.e., the accuracy of the model can be ensured. In additional, the stocks of these three companies are also on a downward trend and will remain floating in a range for some time in the future, roughly the same as the previous analysis of the results using the CNN model.

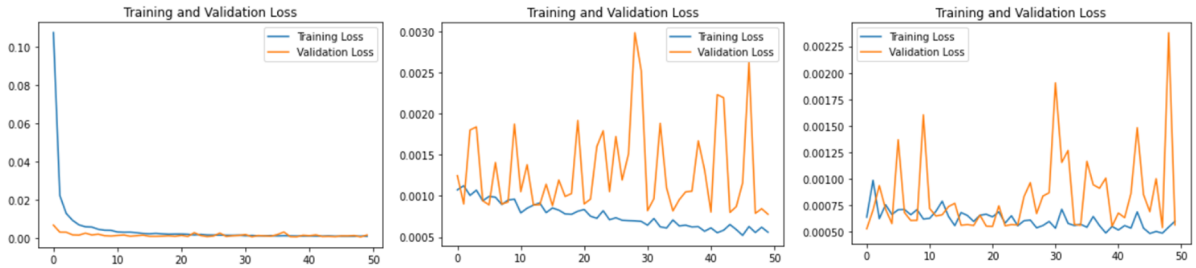


**Fig. 7** the loss function results for Apple, Google and Amazon.

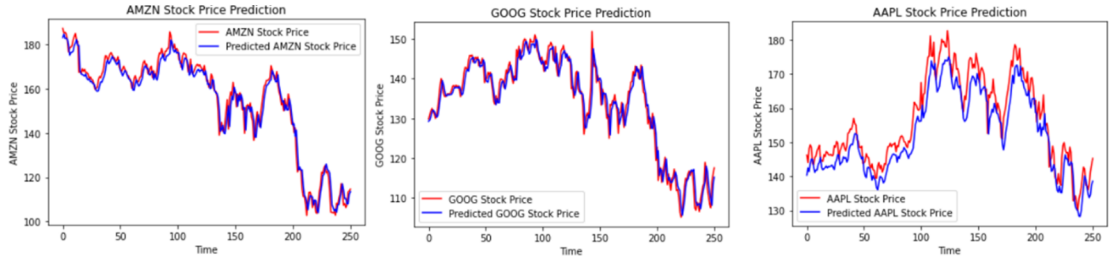


**Fig. 8** RNN model stock price prediction

However, these models have certain research limitations. The results show that it is feasible to process time series data produces predicted values based on CNN but there are somewhat different from the true values. In this case, there are many factors that affect stock price volatility and predicting future stock prices from historical stock prices alone is not sufficient, which is also in line with the weak efficient market hypothesis that technical analysis fails. On the other hand, the network structure used in this experiment is relatively simple and a more powerful network for the regression prediction task would have yielded better results. Regarding to the prediction basde on RNN model, single step prediction must be used. It has the drawback of being inaccurate at time t, which cannot use the sequence information at t+1 and after. Theoretically, the RNN model can use the information of any long sequence. Whereas, it can remember a limited length in practice, which will lead to gradient explosion or gradient disappearance after a certain time. Therefore, a combination approach is more suitable, e.g., in conjunction with an LSTM model, to attempt to address this issue.

## 4. Summary

In conclusion, this paper proposed a CNN model to predict the price of Apple, Google, and Amazon. Primarily, the origin and basic structure of the CNN model are introduced. Based on this model, one proposes the way to make stock predictions, and then makes time-series stock price predictions for three technology companies, Apple, Google, and Amazon, and produces results. Compared with historical stock price data, the stocks of the three companies are trending down, but Apple and Google are down more, while Amazon is down less, predicting a future a period of time where the share price will rise slightly and float between $100-$125 per share. The RNN model was then used to compare with the results of CNN, where both models ending up predicting roughly the same result. However, there are limitations to each model, for example the CNN model used in this experiment has a simpler network structure and a more powerful network for the regression prediction task would have given better results. In the future, this study provides insight into stock market change trend prediction, which can offer a guideline for real-life stock prediction.

# References

[1] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana and Shahab S. Deep Learning for Stock Market Prediction. Entropy, 2020, 22(8): 840.

[2] C. Lee, J. Lee, and C. Lee. Stock prices and the efficient market hypothesis: Evidence from a panel stationary test with structural breaks, Japan and the World Economy, 2010, 22(1): 49-58.

[3] Faten Subhi Alzazah and X Cheng. Recent Advances in Stock Market Prediction Using Text Mining: A Survey. E-Business-Higher Education and Intelligence Applications, 2020.

[4] Rakhi Batra; Sher Muhammad Daudpota. Integrating StockTwits with sentiment analysis for better prediction of stock price movement. 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). IEEE, 2018: 1-5.

[5] Kriti Pawar, Raj Srujan Jalem and Vivek Tiwari. Stock Market Price Prediction Using LSTM RNN. Emerging trends in expert applications and security. Springer, Singapore, 2019: 493-503.

[6] T S Borkar, and L J Karam. DeepCorrect: Correcting DNN models against image distortions. IEEE Transactions on Image Processing, 2019, 28(12): 6022-6034.

[7] J. Wu. Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 2017, 5(23): 495.

[8] Yahoo Finance-Apple Inc. Retrived from: https://finance.yahoo.com/quote/AAPL/history?p=AAPL.

[9] Yahoo Finance-Alphabet Inc. Retrived from:https://finance.yahoo.com/quote/GOOG/history?p=GOOG.

[10] Yahoo Finance-Amazon Inc. Retrived from:https://finance.yahoo.com/quote/AMZN/history?p=AMZN

[11] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction to Statistical Learning, New York: springer, 2013.