A Deep Learning CNN based Approach to the Problem of Crypto-Currency Prediction

Sajan Kumar Kar

SRM Institute of Science and Technology, KTR, Chennai

Corresponding author: Sajan Kumar Kar, Email: sajan.kumar.kar@gmail.com

After the invention of Bitcoin by a man named Satoshi Nakamoto along with other blockchain-based person-to-person payment systems, the cryptocurrency market has instantly gained popularity. Because of this, that is, the volatility of the various cryptocurrency prices. This attracts much attention from both the investors and the researchers. The task of forecasting the prices of crypto-currencies because of the static prices and the arbitrary effects in the market is quite challenging. Cryptocurrency price forecasting models that are available now mainly focus on analysing extrinsic factors, like macro-financial indicators, data linked to the blockchain, and data from social media – with the goal of enhancing the prediction accuracy. However, the intrinsic noise present in the raw data, caused by market and political conditions worldwide, is complex to interpret. In our research we propose a multiple input convolutional neural network model, specifically a convolutional neural network model for the prediction of future cryptocurrency price. Generally, RNNs and LSTMs are used for problems dealing with timeseries data. We used the concept of residual networks on 1-Dimensional convolutional networks to solve the problem of predicting the price of Bitcoin, the most popular cryptocurrency out there at the moment. Furthermore, we conduct additional experiments on ether, the cryptocurrency of Ethereum to further confirm that even CNNs can work equally well, if not better in comparison to the widely used LSTM neural network models.

Keywords: Deep Learning, Cryptocurrency price prediction, CNN.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 1123–1136. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-106

1 Introduction

Cryptocurrency is a type of digital currency or asset, which utilizes blockchain technology and cryptographic functions to gain decentralization and immutability. It is a digital form of money that can be used to buy goods and services via an online ledger with strong cryptography to secure online transactions. A cryptocurrency is a digital exchange medium, where strong and secure cryptographic algorithms, like Secure Hash Algorithm 2 (SHA-2) and Message Digest 5 (MD5), are used to secure the transaction records. Apart from making it secure, it makes the transactions traceable, immutable, and transparent.

Because of these properties, cryptocurrencies, especially bitcoin, have gained immense popularity in almost all the sectors, and mostly in the financial sector. Though cryptocurrencies are in still in the process of getting recognition from the governing bodies, the uncertainty in their prices cause the risks in the investments to rise substantially. As the value of these cryptocurrencies varies every day, it would be very interesting for investors to forecast the Bitcoin value but at the same time making it difficult to predict.

Predicting Cryptocurrency prices can provide a lending hand to investors who indulge in cryptocurrency for making good investment decisions in order to acquire profits of higher margins, while it can also support decision-making and financial researchers for studying crypto-currency markets behaviour. Cryptocurrency price prediction is like a common type of time series prediction problem, like the stock price prediction. Traditional time series prediction methods such as the well-known ARIMA model, have been applied for cryptocurrency price and movement prediction. However, these earlier models were unable to capture non-linear complex patterns of prediction problems in contrast to Deep Learning algorithms which procure better results when applied on the problem including timeseries data.

In this work, a novel Deep Neural Network model consisting of 1-D CNNs with residual network is proposed for Cryptocurrency price prediction. In the study, not only the influence of transaction information such as historical price of Crypto is considered on the closing price of the currency, but also external influences such as macroeconomic factors and investors' attentions are introduced.

The remaining part of the paper is organised as follows. Section II briefs the relevant literature. Section III details the methodology about the proposed work highlighting on the building blocks used in solving the problem. Section IV shows our implementation while on Section V we provide a glimpse of the code. Finally on Section VI we show the results and conclude the report.

2 Literature Survey

Salim Lahmiri et al. [1] first applied DNN to predict cryptocurrency price, and found that the predictive accuracy accrued by LSTM neural networks is greater than that achieved by GRNNs which stand for Generalized Regression Neural Networks [2]. LSTMs have the capability to effectively hold both short- and long-term dependencies in the data. Thus, they have worked significantly well in predicting bitcoin prices. Several research grads and scientists have procured relatively better results with the LSTM networks when dealing with the task of forecasting crypto prices [3].

Patel et al. [4] suggested a hybrid approach to cryptocurrency prediction, in which he used Litecoin and Monero crypto-currencies. The model combines LSTM and GRU layers to form the hybrid RNN architecture. The Litecoin data was had daily values from 24 August 2016-23 February 2020 and the Monero data had values from 30 January 2015-23 February 2020. It consisted of the average price, OHLC, and the volume of trades. The results illustrated that the proposed hybrid model outperformed traditionally employed LSTM networks, thus, showing some promising results.

Livieris et al. [5] proposed a hybrid CNN and LSTM model in which they used 3 consecutive years of cryptocurrency data, i.e. from 1 January 2017 to 31 October 2020, from the three leading crypto-currencies at that time, which were, Bitcoin (BTC), Etherium (ETH), and. This showed the utility of skip connections in residual networks when used with timeseries data. Their experiments showed that this proposed model, a hybrid of CNN and LSTM, exceeds traditionally used LSTM networks exhibiting splendid results.

Another inventive discovery in this field was shown by Vasily Derbentsev [6] in which they described the building of the short term forecasting or prediction model of cryptocurrencies' prices using a machine learning approach.In "Forecasting cryptocurrency prices time series using machine learning approach", from standard regression tree models and time series data, a modified BART- Binary Auto Regressive Tree model was developed. - BART is a hybrid method that combines classification and regression trees.

Mahdiye et. al [7] in their research predicted the direction of trend shift in cryptocurrency. Quitong Guo et. al [8] used a hybrid LSTM and CNN approach to predict the price of Bitcoin in a Multi-scale residual framework. For predicting high frequency trend in crypto exchange rated Monsalve and others [9] used CNNs. Pantelis used data mining and text mining techniques in his research to predict Bitcoin prices. [10]

Ji Suhwan and others [11] during their research did ”A Comparative Study of Bitcoin Price Prediction Using Deep Learning” in which they used Deep Neural Networks, LSTM, CNN, Residual network models individually and combined them to compare as well. The superficial result was as expected, the LSTM networks performed a little better but they also found out that DNN based models best predicted the ups and downs in prices.

There was also a multivariate prediction on cryptocurrency done by Seng Hansun and others [12] who used multivariate prediction approach and used three different RNNs- LSTM, Bi-directional LSTMs and GRU on 5 major cryptocurrencies-Bitcoin, Etherum, Cardano, Tether, and Binance. Their research resulted in the conclusion that Bi-LSTM and GRU resulted in similar accuracy but with regards to execution time, GRU and LSTM depicted similar results. Another exclusive research on GRUs for Bitcoin price prediction [13] resulted that a GRU model with recurrent dropout performs better than the existing models.

3 Methodology

In current section, we explain the details of our proposed work regarding cryptocurrency price prediction using CNN and the basic blocks as briefed earlier.

3.1 Convolutional Neural Network

A CNN is a network of convolutional layers. These convolutional layers are the most integral parts of the CNN that consist of learnable filters called kernels. Each filter is quantifiably a vector of small size (e.g., 2x2 for a 2D BnW image, or 1x4 for a 1D vector). The convolution process involves passing the filter through the image matrices one position at a time with 1 or more strides. The filter is set at a position, then its pixel values are multiplied element-wise with the values of the image matrix. The matrix produced by this is termed as a feature map which are stacked in order to go to act as input for the next layer. This convolution action makes the network learn patterns or features like edges.

Convolution is a mathematical operation where you ”summarize” a tensor or a matrix or a vector into a smaller one. If your input matrix is one dimensional then you summarize along that one dimension, and if a tensor has n dimensions, then you could summarize along all n dimensions. For instance, in Figure 1, consider a 5x5 image (green) which only has values 0 or 1 and a 3x3 filter (yellow):

The convolution operation of the 3x3 filter across the 5x5 image would look something like in Figure 2:

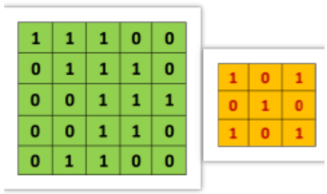

Figure 1: **Convolution matrix and filter** (Rekha S, 2018) [16]

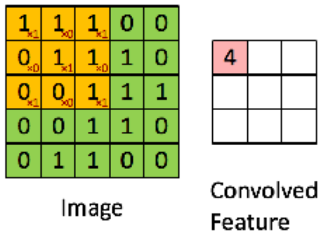

Figure 2: **Convolution operation** (Rekha S, 2018) [16]

The 3x3 filter moves across the image 1 pixel at a time. The filter values are first multiplied elementwise with the pixels of the image at the corresponding position and then the resulting values are added to get the final output which then becomes a single element of the feature map. Different filter values will lead to different output matrices otherwise known as feature maps.

3.2 LSTM

LSTM is a type of RNN which solves the problem of gradient disappearance and gradient explosion that is the major handicap of RNN. It stands for Long Short-Term Memory. LSTMs use a unique gate control mechanism, as shown in Figure 3, which solves the problem of long-term dependency.

3.3 Data

A real-time cryptocurrency data can be found from various sites like Kaggle, Quandl and other website which provide APIs to scrape the data. For our work data has been collected from the website of Poloniex, which is a major crypto asset exchange company. It contains bitcoin data with minute-to-minute updates of OHLC data(Open, High, Low, Close).

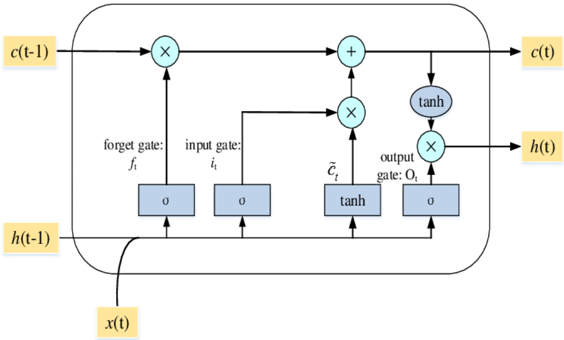

Figure 3: **The basic structure of a LSTM unit** (Tomar A, Ghosh B, 2020) [17]

3.4 Proposed model components

The model will be a combination of 1D Convolutional layers, dropout layers and the activation function used will be ReLU or its variant Leaky ReLU. ReLU stands for Rectified Linear Unit. It is a piecewise linear function defined as shown in Figure **??**:

Figure 4: **ReLU** (own photo)

The ReLU activation function has certain problems. The major being that few of the slopes or gradients can be frail during the training and can, thus, easily die. This leads to an update in the neuron weight that makes it problematic to activate ever. In simple words, the ReLu activation function can result in dead neurons. To solve this we use its variant Leaky ReLU,(**ng**) which provides a small slope for negative values, instead of cutting it to zero. For instance, leaky ReLU may has(say) alpha=0.01, so the equation becomes as shown in Figure 5:

$$R(z) = \left\{ \begin{array}{ll} z & z > 0 \\ \alpha z & z <= 0 \end{array} \right\}$$

Figure 5: **Leaky ReLU** (own photo)

The graph for both the activation functions looks like the one in Figure 6:

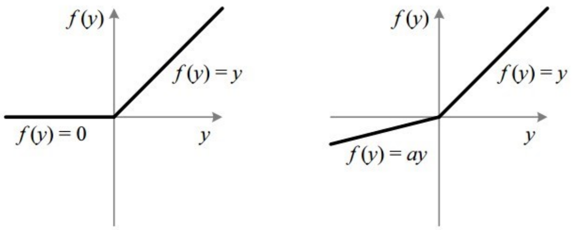

Figure 6: **Left- ReLU, Right- Leaky ReLU** (Sharma S, 2017)] [18]

4 System Architecture and Design

The data collected from the source is parsed in order to be sent to the model for training and prediction. A glimpse of the data is shown in Figure 7

Out[16]:

	Close	Timestamp	High	Low	Open
0	1.75	1439014500	0.33	1.61	0.33
1	1.85	1439014800	1.85	1.85	1.85
2	1.85	1439015100	1.85	1.85	1.85
3	1.85	1439015400	1.85	1.85	1.85
4	1.85	1439015700	1.85	1.85	1.85

Figure 7: **Dataset sample** (own photo)

Now to note, this is just the data for Bitcoin. Further data has also been collected for Ethereum and Binance Coin. That data is also in the similar OHLC format.

Various pre-processing is done on the data like scaling and regularization. Regularization is a regression technique used to tune the cost function. It adds an additional penalty term in the error function which reduces the excessively fluctuation in the function in order for the coefficients to not take extreme values. In other words, this technique discourages learning a more complex model in order to prevent the risk of overfitting. The model architecture is shown below in Figure 8

To measure the accuracy of the predicted values we use MSE metric which stands for Mean Squared Loss. MSE estimates how close a regression line is to a set of points. It calculates the distance from the data points to the regression line (these distances hold more weight to larger differences). Then the MSE evaluates the average of the set of errors, the formula for which is shown in Figure ??.

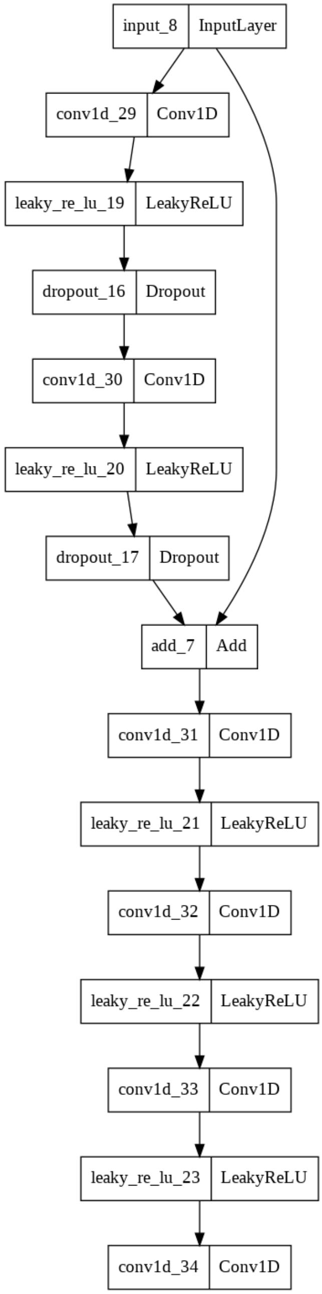

Figure 8: **Model Architecture** (own photo)

Figure 9: **Mean Square Error** (Abhigyan, 2020) [19]

where, n = the number of items, sigma = summation notation, y = the original or observed y-value, y-hat = the y-value from regression.

5 Coding and Testing

We use the API provided by the Crypto exchange Poloniex to collect the data. A simple call to the API and parsing it with json.loads() provides us with the data which can be stored in dataframes included in pandas library in python using pd.DataFrame(). We parse it using a PastSampler class. Then we use MinMaxScaler() to scale and normalize the data.

Now that the data is gathered and the pre-processing is in place, we proceed to building the model. Using the Keras Model() API, we code the architecture of the model building residual networks [15]. It looks something like:

```
model= Model(inputs=layer_in, outputs=layer)
#The code for the layers look for instance,
conv1= (Conv1D(input_shape=(step_size, nb_features), padding='same',
filters=8, kernel_size=8))(layer_in)
leakyRelu1= LeakyReLU()(conv1)
dropout1= Dropout(0.5)(leakyRelu1)
```

Then, after mentioning the necessary hyperparameters and parameters, we compile the model and run the model using model.fit().

Once the model has been trained it is time for predicting the prices and displaying it in graphs which is achieved using matplotlib library. The accuracy of the model is measured using MSE which stands for mean squared error.

6 Results

We implemented the approach using our model which has a single residual unit and also using a single LSTM unit in order to compare and show that CNNs with residual units can match the performance of the state-of-the-art LSTMs used for time series prediction. The results we got were interesting to say the least. Below is the output prediction graph of both approaches for the different Cryptocurrencies.

6.1 BITCOIN

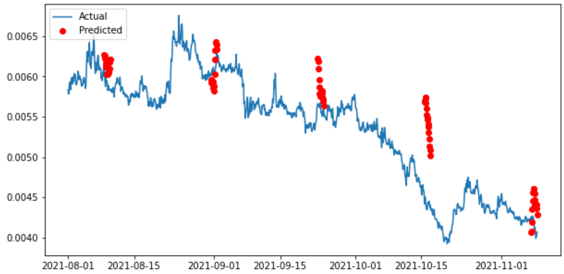

Figure 10: **CNN Output Graph(BTC)** (own photo)

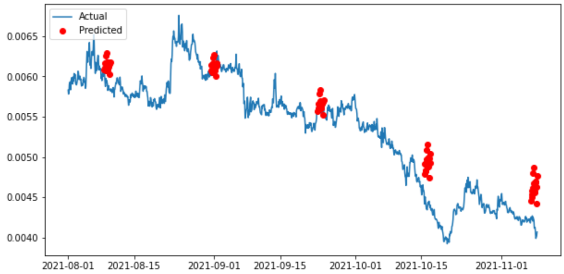

Figure 11: **LSTM Output Graph(BTC)** (own photo)

Both the graphs in Figures 10 and 11 show the actual prices (blue line chart) and the predicted prices at intervals at that point of time(red dots). There is not a significant difference but we can see that at times where there is a sudden change is price, the LSTM network is able to capture it a little better than the residual network. Look at the prediction for October 15 for example. The prediction by the LSTM network is a little closer to the actual value. Hence, in this case of bitcoin the LSTM network can be considered to perform better than our proposed CNN model with residual networks.

6.2 ETHEREUM

This is where the results get interesting. Earlier in bitcoin we saw the LSTM out-performing residual network by a narrow margin. But here, in Ethereum, the results clearly show that the residual model has performed better than the LSTM model. The output plots are shown below in Figures 12 and 13.

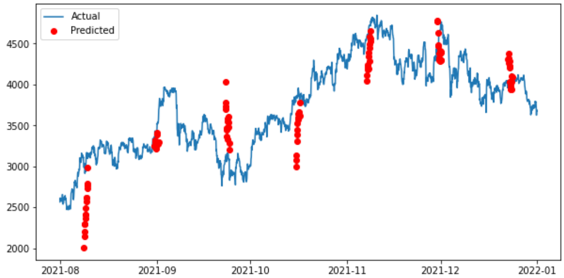

Figure 12: **CNN Output Graph(ETH)** (own photo)

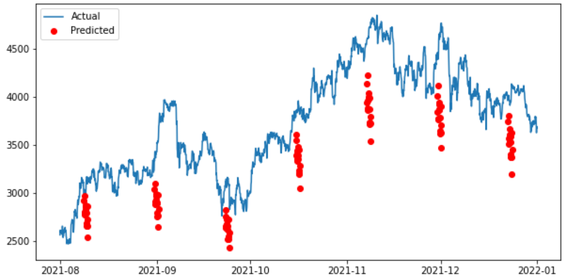

Figure 13: **LSTM Output Graph(ETH)** (own photo)

The predicted values of the CNN residual model are closer to the ground truth values as is evident from the plot. The LSTM model is able to capture the trend but the values it predicts looks as if it is at a certain offset from the actual value.

6.3 BINANCE COIN

The result for binance coin is kind of similar to that for ethereum. It is noticed that the LSTM model captures the trend but isn't able to aptly predict the actual prices and the predictions have quite a large offset. The resultant plots are shown below in Figures 14 and 15.

The CNN residual model however doesn't work as good as it worked for ethereum. However, still, it works better than the LSTM model here. It is noticed that, while the residual model isn't able to appropriately predict in uptrend and downtrend, in sideways trend the residual model works pretty well.

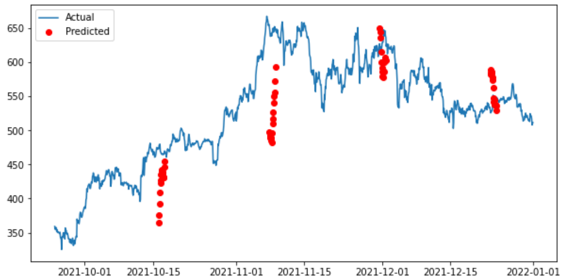

Figure 14: **CNN Output Graph(BNB)** (own photo)

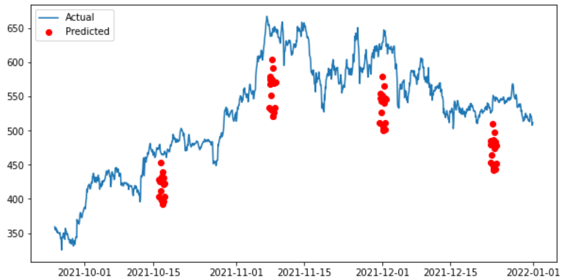

Figure 15: **LSTM Output Graph(BNB)** (own photo)

7 Conclusion

The model has been built using one-dimensional Convolutional Neural Networks. It gives a reasonable degree of accuracy, in comparison with LSTMs which are suited for Time-Series Data. However, as demonstrated in the model, the concept of residual networks when used with 1-dimensional convolutional neural networks works just as well, if not better, as the state-of-the-art LSTM and other hybrid models, with these being less computationally expensive and hence better.

All in all, we have discussed a solution or rather an alternative approach to predicting time series data using a model which is less computationally expensive but performs just as good, if not better. It doesn't come out to be the clear winner but it can be said that using this approach is a really good and possible alternative.

References

[1] Lahmiri, Salim and Bekiros, Stelios. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons and Fractals. 118. 35-40. 10.1016/j.chaos.2018.11.014.

[2] D. F. Specht et al. (1991) A general regression neural network IEEE transactions on neural networks, vol. 2, no. 6, pp. 568–576, 1991.

[3] N. Uras, L. Marchesi, M. Marchesi, and R. Tonelli. (2020) Forecasting bitcoin closing price series using linear regression and neural networks models arXiv preprint arXiv:2001.01127, 2020.

[4] Patel, M.M.; Tanwar, S.; Gupta, R.; Kumar, N. (2020) A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions J. Inf. Secur. Appl. 2020, 55, 102583. [CrossRef].

[5] Livieris, I.E.; Kiriakidou, N.; Stavroyiannis, S.; Pintelas, P. (2021) An Advanced CNN-LSTM Model for Cryptocurrency Forecasting. IElectronics 2021, 10, 287. https://doi.org/10.3390/electronics 10030287.

[6] Derbentsev, Vasily and Datsenko, Natalia and Stepanenko, Olga and Bezkorovainyi, Vitaly. (2019) Forecasting cryptocurrency prices time series using machine learning approach. SHS Web of Conferences. 65. 02001.

[7] Cherati, Mahdiye and Haeri, Abdorrahman and Ghannadpour, Seyed. (2021). FCryptocurrency direction forecasting using deep learning algorithms. Journal of Statistical Computation and Simulation. 91. 1-15.

[8] Guo, Qiutong and Lei, Shun and Ye, Qing and Fang, Zhiyang. (2021). MRC-LSTM: A Hybrid Approach of Multi-scale Residual CNN and LSTM to Predict Bitcoin Price. 2021 International Joint Conference on Neural Networks (IJCNN), 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9534453.

[9] Alonso Monsalve, Saúl and Suárez-Cetrulo, Andrés and Cervantes, Alejandro and Quintana, David. (2020) Convolution on Neural Networks for

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 1123–1136. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-106

High-Frequency Trend Prediction of Cryptocurrency Exchange Rates Using Technical Indicators. Expert Systems with Applications. 149. 113250. 10.1016/j.eswa.2020.113250.

[10] Linardatos, Pantelis and Kotsiantis, Sotiris. (2020). Bitcoin Price Prediction Combining Data and Text Mining. Intelligent Technologies and Robotics https://doi.org/10.1007/978-981-15-1918-5-3

[11] Ji, Suhwan and Kim, Jongmin and Im, Hyeonseung. (2019) A Comparative Study of Bitcoin Price Prediction Using Deep Learning. Mathematics. 7. 898. 10.3390/math7100898.

[12] Hansun, Seng and Wicaksana, Arya and Khaliq, Abdul. (2022). Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches. Journal of Big Data. 9. 10.1186/s40537-022-00601-7.

[13] Dutta, Aniruddha and Kumar, Saket and Basu, Meheli. (2020). A Gated Recurrent Unit Approach to Bitcoin Price Prediction. Journal of Risk and Financial Management. 13. 23. 10.3390/jrfm13020023.

[14] Maas, A.L., Hannun, A.Y. and Ng, A.Y. (2013) Rectifier Nonlinearities Improve Neural Network Acoustic Models. Proceedings of the 30th International Conference on Machine Learning, Vol. 28, 3.

[15] He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian. (2016) Deep Residual Learning for Image Recognition. 770-778. 10.1109/CVPR.2016.90.

[16] Rekha, Shashi. (2018). Deep Indian Delicacy: Classification of Indian Food Images using Convolutional Neural Networks.

[17] Tomar, Ayushi and Ghosh, Bikramaditya and Manjunath, Chinthakunta and Addapalli, V and Krishna,. (2020). Employing Deep Learning In Intraday Stock Trading. 10.1109/ICRCICN50933.2020.9296174..

[18] Sharma, Sagar. (2017). Activation Functions in Neural Networks. https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[19] Abhigyan. (2020). Calculating Accuracy of an ML Model. https://medium.com/analytics-vidhya/calculating-accuracy-of-an-ml-model-8ae7894802e