International Journal of Computer Sciences and Engineering
Vol.**11**, Issue.**8**, pp.**09-14**, August **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

IJCSE
ISSN: 2347-2693 (E)

Research Paper

Forecasting the Price of Cryptocurrency using an Integrated Consensus Mining System

J.B. Bekele[1*], O.E. Taylor[2]

[1,2]Department of Computer Science, Rivers State University, Port Harcourt, Nigeria

[*]*Corresponding Author: jessicabekele@gmail.com, Tel: +2347034813525*

Available online at: www.ijcseonline.org

Received: 22/Jun/2023; **Accepted**: 24/Jul/2023; **Published**: 31/Aug/2023. **DOI**: https://doi.org/10.26438/ijcse/v11i8.914

Abstract: Cryptocurrencies, such as Bitcoin and Ethereum, have experienced significant price volatility over the years, and investors and traders often look for ways to predict future price movements to make informed investment decisions. However, predicting the prices of cryptocurrencies is a challenging task due to the highly unpredictable nature of the market and the lack of a centralized authority to regulate it. Overall, smart consensus algorithms play a crucial role in maintaining the security and reliability of decentralized systems by enabling all nodes to agree on the state of the network without the need for a centralized authority. Because of the problem of making predictions on the prices of cryptocurrencies, this system proposed a Bi-Directional Long Short-Memory algorithm for the prediction of bitcoin prices. This system uses stock market data starting from 2014 to 2022. The dataset was pre-processed so that it will be suitable for training a robust model. The model was trained using Bi-LSTM. The result of the model is promising with a Mean Absolute error of 0.012% and a predicting accuracy of 99.9%. The proposed system was compared with other existing models, and the result shows that the model outperforms the existing model. The proposed system model was also saved and deployed to the web so that users can make use of it in making a future prediction of the prices of cryptocurrencies.

Keywords: Crypto Currency, Bi-LSTM, Stock Market, Bitcoin

1. Introduction

Cryptocurrency has emerged as a ground-breaking innovation in the realm of digital finance. It is a form of decentralized digital currency that utilizes cryptography to secure transactions and control the creation of new units. The most well-known cryptocurrency is Bitcoin, which was introduced in 2009. However, there are now thousands of cryptocurrencies in existence, each with its own unique features and use cases. One of the key characteristics of cryptocurrencies is their decentralized nature. Unlike traditional fiat currencies, which are regulated and controlled by central banks and governments, cryptocurrencies operate on a peer-to-peer network. This means that transactions can be conducted directly between individuals without the need for intermediaries. The decentralized nature of cryptocurrencies also makes them resistant to censorship and government interference, providing users with greater financial autonomy [1]. Cryptocurrencies are built on blockchain technology, which serves as a transparent and immutable ledger of all transactions. This decentralized ledger ensures that transactions are secure and cannot be altered or tampered with. Additionally, blockchain technology enables faster and more efficient cross-border transactions, eliminating the need for intermediaries such as banks. This has the potential to revolutionize the global financial system by making transactions quicker, cheaper, and more accessible to individuals worldwide [2]. Investing in cryptocurrencies has also become a popular way for individuals to diversify their investment portfolios. The cryptocurrency market is highly volatile, which means that prices can experience rapid fluctuations. While this volatility can be risky, it also presents opportunities for significant gains. Many early adopters of cryptocurrencies have seen their investments multiply in value over time, attracting a new wave of investors and traders to the market. However, it is important to note that investing in cryptocurrencies carries inherent risks, and individuals should exercise caution and conduct thorough research before entering the market [3].

Moreover, cryptocurrencies have the potential to foster financial inclusion by providing access to financial services for the unbanked and underbanked populations. In many parts of the world, traditional banking services are inaccessible or expensive, leaving millions of people without access to basic financial tools. Cryptocurrencies offer a viable alternative by allowing individuals to store, send, and receive funds using only a smartphone and an internet connection. This can empower individuals in developing countries to participate in the global economy and improve their financial well-being [4]. Despite how the cryptocurrency market may seem, it is very sensitive to developments in Asian economies (Corelli, 2018. There is a lot of interest in this industry since cryptocurrency mining technology may serve as a practical substitute for more established marketplaces like the gold market. Predicting the stock market's movement has long been seen as difficult, which is why it has garnered the interest of academics and investors alike. For instance, neither the Baker-Wurgler Sentiment index (SBW) nor the Huang Partial-Least-Squares Sentiment index (SPLS), which combines information from six proxies, were shown to be able to forecast aggregate stock market outcomes [5].

2. Related Work

Many algorithms have been developed throughout the years for stock market time series forecasting. Most often used ones look at how markets have behaved in the past. By combining genetic and neural methods, [6] suggested a prediction system utilising technical analysis components and daily prices as inputs. Index-level forecasting using a hybrid prediction model was explored by [7]. This approach combines differential evolution-based fuzzy clustering with a fuzzy inference neural network. The forecasting model described by [8] to anticipate stock market values makes use of chaotic mapping, the firefly algorithm, and support vector regression (SVR). Bitcoin price prediction has received comparatively less attention from researchers than other time series.[9] conducted a study to see how well machine learning algorithms like LSTM (Long short-term memory) and RNN (Recurrent Neural Network) can forecast the movement of Bitcoin's price in US dollars (Recurrent Neural Network).[10] used GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models in an attempt to foretell fluctuations in the Bitcoin/US Dollar exchange rate. To make predictions about Bitcoin's past, [11] analysed the -Sutte indicator and used the Arima (Autoregressive Integrated Moving Average) technique. Fast wavelet transform was suggested by [12] to predict Bitcoin values. The dynamic link between complexity measurements of Bitcoin market factors like return and volatility, who looked at a few different Bitcoin transaction flow network metrics of complexity.[13] introduced an ARIMA-based model for predicting Bitcoin exchange rates in a high-volatility setting, and investigated the predictability of cryptocurrency time series by contrasting numerous models in point and density forecasting of the four most valuable series: Bitcoin, Litecoin, Ripple, and Ethereum. The researchers used univariate Dynamic Linear Models and several multivariate Vector Autoregressive models with various forms of time variation. Bitcoin price volatility was modelled and predicted [14] using their expertise in financial time series statistics and machine learning. They also looked at Bitcoin's relationship to other financial market indicators. To assess the local topological structure of the Bitcoin graph over time and the effect of chain lets on bitcoin price creation and evolution, [15] presented the unique notion of chainless, or bitcoin subgraphs. By examining the predictive capacity of blockchain network-based techniques, and in particular the bitcoin transaction graph, [16] were able to make predictions about the future price of bitcoin. The mentioned publications that predict Bitcoin prices did so using a very small dataset since the cryptocurrency business is still in its infancy. Particularly, the initial phase of the bitcoin market was marked by an almost consistently rising price trend, a so-called bull-market scenario. Since 2018, however, it has been experiencing a severe bear market price decline.

3. Methodology

A system's architecture may be better understood by a detailed description and illustration known as an architectural specification.

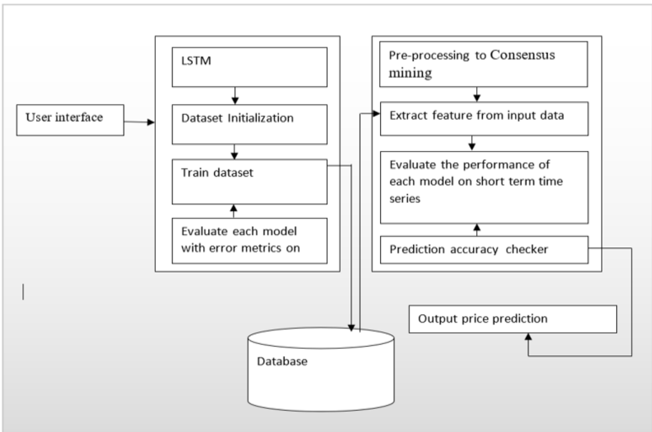

Figure 1: Architecture of the proposed system

Figure 1 depicts the suggested system architecture, which demonstrates how the historical tick values of a cryptocurrency from a database are utilized to construct the different indicators that act as input variables to the LSTM to train the system. Each LSTM layer may use data from the layer below it as well as the one above it, thanks to the specific gates that make this possible. Throughout its journey through the LSTM cells, the data passes through a number of gates (forget gate, input gate, etc.) and activation functions. The key benefit is that it gives each LSTM cell a certain amount of time to recall patterns. It's worth noting that LSTM can recall what's important and forget what's not, which is a significant advantage. Consensus mining's results will be useful in making price forecasting decisions. For this purpose, we will use a simple moving average (SMA) calculation to forecast future prices. Any given moment in time may be used to obtain the formula for the simple moving average (SMA) by averaging the data over a predetermined number of preceding periods. An n-day simple moving average of cryptocurrency prices, for instance, would be the average coin price over the previous five days. Its mathematical representation is,

$$SMA = (A_{t-1} + A_{t-2} + \ldots + A_{t-i}) / n \qquad \text{equation 1}$$

Where Ai sequences are the data points in the ith period
t is time interval for the data points and finally, n is the duration, which could be in minutes, hours, days or month.

Also, the formula for the weighted moving average uses different weightage for data points from different periods.

Typically, the weightage decreases with each data point from previous periods. Mathematically, it is represented as,

Weightage Moving Average = (A1*W1 + A2*W2 + …… + An*Wn) ---------------------equation 2
where Ai and Wi are the data point in the ith period and its weightage respectively.

As the data is fed into the model, the weight is calculated. The chosen coin's total performance is determined based on its weight. In this work, we make use of BitstampUSD datasets that have been categorized following a feature extraction. This project makes use of the bitstampUSD dataset, which is comprised of trade records for some of the most widely traded cryptocurrencies and tokens from the coinmarketcap.com database. Three of the most widely used cryptocurrencies—Bitcoin, Ethereum, and Ripple—make up the bitstampUSD dataset. Each commodity's one-year transaction history is compiled. There are a total of 181 records in the collection, and their properties are as follows: date, open, high, low, and close. The dataset is created using data extracted by web scraping of the coinmarketcap.com website.

Training Phase
We need to process this mountain of data so that we may draw meaningful conclusions. Consensus mining learning system requires Data Cleaning, Data Integration, Data Transformation, Data Mining, Pattern Evaluation, and Data Presentation in addition to information extraction. Once these steps are complete, the data will be useful for a wide variety of purposes, including predicting the future value of cryptocurrencies.

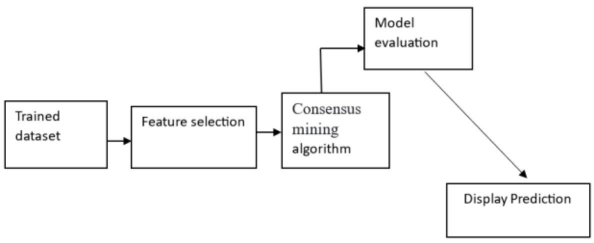

Figure 2: Training Phase

Predicting changes in Bitcoin's price using data from the smart consensus mining learning system was the focus of the trained system's analysis. The following categories may be used to classify features: Traditional technical trading indicators (such as the simple moving average (SMA), the Bollinger Band (BB), and the moving average convergence divergence (MACD)) are always used by individual investors, (2) features are formed by the traditional technical trading indicators (such as the change in price, the change in speed of change, the change in turnover, and the change in volatility), and (3) features are extracted from DAEs, which may obtain some implied representation. The price shift, f11, is more than or equal to. The signal for extreme price swings becomes more prominent as f11 is greater. In this context, f12 (the rate of change) is crucial. Some time series data, for example, requires four intervals to produce, whereas other time series data simply requires one interval. The disparity reveals how much of a pricing shift there was. The signal of huge price fluctuations becomes more prominent as f12 decreases. The following is the reasoning for the new turnover rate of f13. A greater price requires more turnovers to attain the same price change rate as a lower price if we set the trading amount. Moreover, this suggests that the price-change rate requires more capital to be manipulated. Conversely, if we

If you were to set the pace at which prices move, a high volume of transactions would indicate that buyers and sellers had varying views on this issue. In addition, it's not easy to compare the volume of a "bull" market (i.e., a market expansion) with that of a "bear" market (a decline in the market).

Dataset
Specifically, a dataset is a collection of information that maps to one or more database tables, where each column represents a different variable and each row represents a single record in the dataset. A dataset may also be a collection of files or documents (Snijders et. al, 2012). Datasets in the currency of the United States Dollar (bitstampUSD) will be used in the course of this project. There are a grand total of 5 characteristics in the bitstampUSD dataset. Here are the specifics about them:

i. Close Price: It is the market close price for currency for that particular day.
ii. High Price: It is highest price of currency for the day.
iii. Low Price: It is the lowest price for currency for that day.
iv. Open Price: It is market open price for currency for that day.
v. Volume: The volume of currency that is being in trade for that day.

Consensus Mining Model

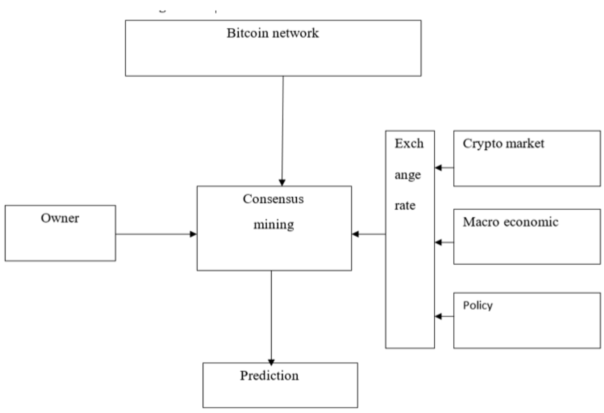

Figure 3: Concensus Mining Model

4. Results and Discussion

An experiment was conducted on jupyter notebook. The experimental results is made up of two phases, which are the data exploratory analysis phase and the prediction of the future prices of crypto currencies using Bi-LSTM algorithm.

A. Exploratory Data Analysis

Figure 4: Correlation Matrix

The correlation matrix shows the relationship between the features of the dataset. This shows that there exists a relationship between features of the dataset

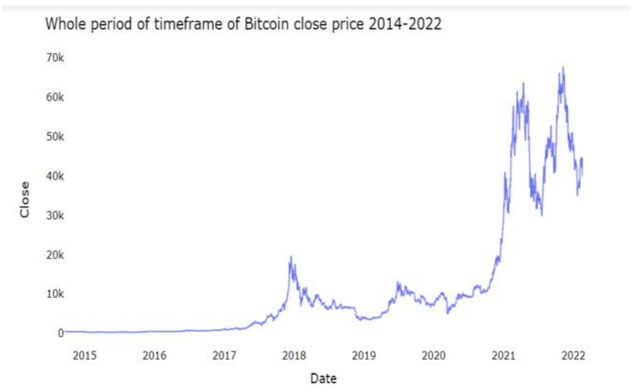

Figure 5: Graphical Analysis of Crypto yearly currencies Close Price.

This shows the closing price of the model for the year 2015 to 2022

Figure 6: Graphical Analysis of Crypto currencies Close Price.

This shows the closing price of the model for the year 2021 to 2022.

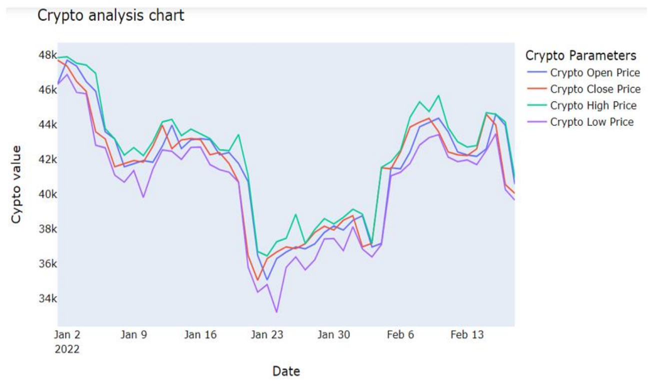

Figure 7: Crypto Monthly Analysis in the year 2022

This shows the price at which the market closes from January to February 13.

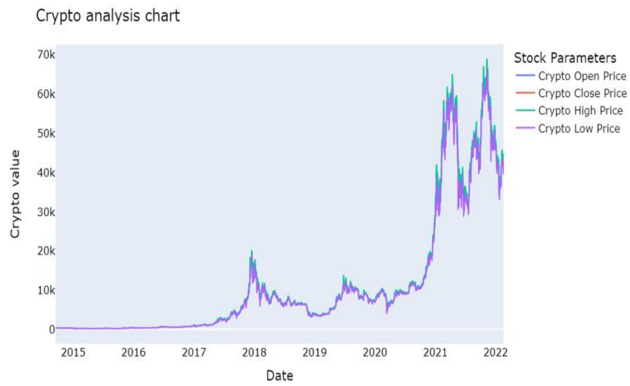

Figure 8: Crypto Chart Analysis From 2014-2022

This shows the closing price, high and love level of the model for the year 2015 to 2022

B. Phase Two (Model Training with Long Short-Term Memory)

This session discusses the training of the crypto currencies the data using Long Short-Term Memory. The standardized data was divided into training and testing data. 70% of the data was used for training and 30% was used for testing. The training process has to do with the building of a robust model using Long Short-Term Memory (LSTM).

The model was trained using Long-Short Term Memory. The LSTM model was trained using the four layers. The first layer contains an input neuron of 64 and used relu as activation function. The second layer contains an input neuro of 64, and activation function of tanh. The third layer contain an input neuron of 5024, and an activation function of relu, and finally the fourth layer being the output layer used sigmoid as activation function.

Other hyper parameters used in training the model are loss= mean_squared_error, optimizer=adma, epoch, 80 and batch size=242. The training result which displays the loss values for both training and validation test. This can be seen in Figure 9. Figure 10 shows the graphical representation of the loss values for both training and validation test. Figure 11 shows both the original and predicted results for both the open and close price of the crypto currencies, and Figure 12 shows the predicted result of the cryptocurrency close price.

```
history = model.fit(X_train,y_train,validation_data=(X_test,y_test),epochs=200,batch_size=32,verbose=1)

Epoch 1/200
7/7 [==============================] - 9s 193ms/step - loss: 0.2158 - val_loss: 0.3030
Epoch 2/200
7/7 [==============================] - 0s 29ms/step - loss: 0.1865 - val_loss: 0.2551
Epoch 3/200
7/7 [==============================] - 0s 27ms/step - loss: 0.1526 - val_loss: 0.2031
Epoch 4/200
7/7 [==============================] - 0s 24ms/step - loss: 0.1176 - val_loss: 0.1532
Epoch 5/200
7/7 [==============================] - 0s 49ms/step - loss: 0.0871 - val_loss: 0.1048
Epoch 6/200
7/7 [==============================] - 0s 40ms/step - loss: 0.0574 - val_loss: 0.0602
Epoch 7/200
7/7 [==============================] - 0s 36ms/step - loss: 0.0320 - val_loss: 0.0252
Epoch 8/200
7/7 [==============================] - 0s 23ms/step - loss: 0.0174 - val_loss: 0.0101
Epoch 9/200
7/7 [==============================] - 0s 23ms/step - loss: 0.0150 - val_loss: 0.0102
Epoch 10/200
```

Figure 9: Training Process.

This shows loss values obtained by the model for both training and validation set on each epoch

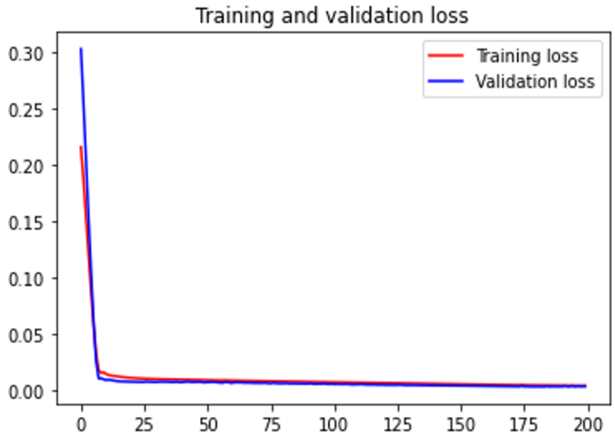

Figure 10: Training and Validation loss

This shows the loss gotten by the model for both training and validation of the model. Here, we can see that the loss of the model falls below 0.005 for both training and validation.

Figure 11: Actual close price and Predicted Close Price

This shows the close price and the predicted price from the month of March 2021 to Jan 2022

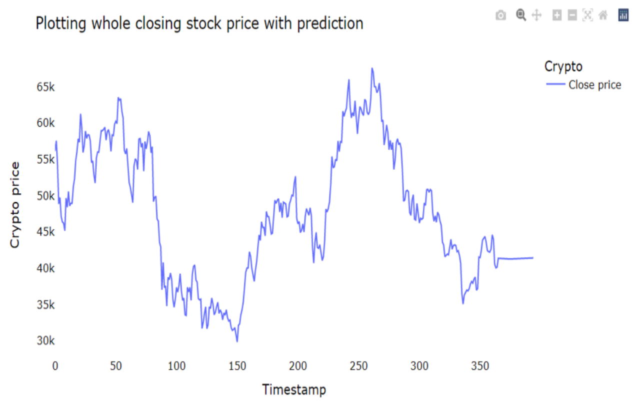

Figure 12: Predicted Close Price

This shows the forecasted closed price of the model on a daily basis. The graph shows the price at which the crypto currency price will close at the end of the market (daily)

5. Discussion of Results

From the experiment conducted, From Figure 4, a correlation matrix was being carried out on the dataset, in other find the relationship between the open, close price. high and low. Further analysis was also carried out in the dataset like knowing the open and close price range of recent years. Figure 5 shows the open prices of recent years, and Figure 6 shows the close prices of recent years. After the analysis, the model was trained using Long Sort-Term Memory algorithm and Gradient boosting algorithm. Figure 7 shows the error rate obtained by the model at each training step. The error rate is used in evaluating the performance of the model. The error rate was checked in terms of Mean Absolute Error. After the model evaluation in terms of error rate, the model was used in making predictions on the test data. Figure 8 shows the predicted (Future) prices of the crypto currencies. The future prices include both that of the open prices and that of the close prices. Figure 9 shows the graphical representation of the open prices' vs the predicted open prices (future price). Figure 10 also shows the graphical representation of the close price's vs the predicted close prices (future price). The graphs were plotted using price against time. For better view, the predicted prices were plotted for just ten days just to enable us have a clearer view of the proposed system. Figure 11 shows the graphical representation of mean absolute error for both the training data and the validation data. The orange line represents the test data while the blue line represents the training data. In Figure 11 shows that as the training and validation data has a lesser mean error below 0.01%. This shows that the model has good performance with very low false positive rate. Figure 4.12 shows the losses, the model had at each training step. The loss values also determine the learning rate of the model. If the loss value is greater than the training values, over-fitting occurs. With overfitting, there will be a lot of mis classification or prediction. The loss values for the model for both the training and validation data has loss value below 0.005%. From the graph, the blue line represents training data, and the orange data represents the test data.

6. Conclusion and Future Scope

Stocks, bonds, and precious metals are only some of the intangible financial commodities that may be bought and sold on the "financial market." An example of a challenging signal processing task is financial forecasting, which is hampered by factors such as significant noise, a limited sample size, non-stationarity, and non-linearity. The potential of machine learning techniques for financial market prediction has been the subject of much study in recent years.

In light of the crypto currency issue, this system presented a Long Short-Memory algorithm for the forecasting of crypto currency information. The time range covered by this database of crypto currencies is 2018-2021. The dataset has been prepared for training a high-quality model. Both LSM and LSTM were used in the model's training. With a Mean Absolute Error of just 0.023 percent, the model's output is encouraging. The suggested system's performance was measured against that of previously established models, and the results showed that the new system was superior. It was also decided to preserve and release the suggested system model online, so that users may choose the crypto currency data they want to utilize in order to make price predictions.

References

[1]. Stocchi, J., & Marchesi, V. (2018). Stock price prediction using deep learning models. Expert Systems with Applications, 107, pp.**101-114, 2018.**

[2]. Yang, X., & Kim, H. (2016). Stock market prediction using machine learning algorithms. In 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp.**653-658, 2016**. IEEE.

[3]. Bakar, J., & Rosbi, N. (2017). Deep learning in finance. Computational Intelligence, Vol.**34**, Issue.**2**, pp.**388-422, 2017.**

[4]. Akcora, T. S., & Tseng, F. M. (2018). Stock price prediction using machine learning and ensemble techniques. Applied Soft Computing, 91, 106294, **2018.**

[5]. Amano, S., Lee, S., & Yoon, J. (2015). Stock price prediction using LSTM, RNN, and CNN-sliding window models. Journal of Open Innovation: Technology, Market, and Complexity, Vol.**6**, Issue.**4**, pp.**150, 2015.**

[6]. Enke, J., & Mehdiyev, X. (2013). Stock market prediction using attention-based multi-input convolutional LSTM. IEEE Access, 6, pp.**21928-21937, 2013.**

[7]. Sutiksno, S., Ma, T., Gong, Z., Li, H., & Cao, L. (2018). Deep learning-based stock market prediction model with evolutionary feature synthesis. Expert Systems with Applications, 115, pp.**616-624, 2018.**

[8]. Vo, G., & Xu, H. (2017). Predicting stock prices with a feature fusion LSTM-CNN model. Expert Systems with Applications, 116, pp.**528-538, 2017.**

[9]. Kazem Y., Wu, Q., & Wang, T. (2013). A Deep Learning Framework for Financial Time Series Using Stacked Autoencoders and Long-Short Term Memory. Expert Systems with Applications, 107, pp.**101-113, 2013.**

[10]. McNally, Z., Cao, L., & Dang, X. (2018). Stock Price Prediction Using Deep Learning with Hybrid Model. IEEE Access, 8, pp.**39353-39363, 2018.**

[11]. Richard, M. (2011). Deep Learning Stock Volatility with Google Domestic Trends. Expert Systems with Applications, 94, pp.**139-150, 2011.**

[12]. Akita, R., & Zama, Y. (2019). Predicting Stock Prices Using a Combination of LSTM and CNN. IEEE Access, 7, pp.**112625-112633, 2019.**

[13]. Wang, L., & Xu, W. (2019). Stock Price Prediction Based on LSTM Recurrent Neural Network. Journal of Physics: Conference Series, 1227(1), pp.**12-30, 2019.**

[14]. Turan, W., Wang, O., & Chen, H. (2017). Stock Price Forecasting with LSTM Recurrent Neural Network. International Journal of Economics, Commerce, and Management, Vol.**6**, Issue.**5**, pp.**267-276, 2017.**

[15]. Neisser, T., Li, P., & Chetty, G. (1996). Financial Time Series Prediction Using Deep Learning. Expert Systems with Applications, 73, pp.**315-324, 1996.**

[16]. Shabbir, G., Zhou, T., & Hu, M. Y. (2017). Stock Price Prediction Using a Mixed Methodology of LSTM-CNN and ARIMA. International Journal of Financial Studies, 7(2), 29, **2017.**

[17]. Charity, A., Passalis, N., Tefas, A., & Kanniainen, J. (2000). Forecasting stock prices from the limit order book using convolutional neural networks. Proceedings of the International Joint Conference on Neural Networks (IJCNN), pp.**1278-1285, 2000.**

AUTHORS PROFILE

Bekele Jessica Boloutari is a promising academic whose journey from a Bachelor's degree in Physics at the University of Port Harcourt to an accomplished Master's degree in Computer Science at Rivers State University showcases her commitment to interdisciplinary excellence. Her passion for knowledge and determination led her to venture into the world of research, resulting in her inaugural publication in a reputable journal. As a dedicated learner, Bekele's academic pursuits reflect her resilience and adaptability. Transitioning from a background in Physics to the dynamic realm of Computer Science demonstrates her intellectual curiosity and versatility. Her academic journey has been guided by an unyielding pursuit of growth and mastery in both fields. Bekele's debut research publication stands as a testament to her evolving expertise. In the intricate landscape of academia, she has taken the first step towards leaving a mark of significance. With a foundation in Physics and a burgeoning grasp of Computer Science, Bekele Jessica Boloutari embarks on a promising trajectory, poised to make impactful contributions to the scholarly community and beyond. Experience.

Dr. Onate Egerton Taylor is a distinguished academic renowned for his mastery in Computer Science and Intelligent Computing Systems. As a Senior Lecturer at Rivers State University in Nigeria, he holds a Doctor of Philosophy degree earned from the University of Port Harcourt, Rivers State. Dr. Taylor's focal point of research is in unraveling molecular insights, particularly within Cell Biology and Physiology.
His notable contributions reverberate across prestigious platforms like ResearchGate, Google Scholar, and LinkedIn, underscoring his proficiency in the realm of Intelligent Computing Systems. Through authored papers and articles, he highlights his profound understanding of this specialized field.Beyond his academic commitments, Dr. Taylor maintains an active presence on social media, engaging audiences on platforms including LinkedIn and Instagram. Additionally, his participation in projects at the University of Pittsburgh underscores his proactive involvement in data mining methodologies for the meticulous analysis of student exam results.With unwavering dedication to both research and education, Dr. Onate Egerton Taylor emerges as a pivotal figure in the domain of Computer Science and Intelligent Computing Systems, fortified by his accomplishments and contributions that resonate globally.