A Deep Learning Approach for Crypto Price Prediction

Sonal Kashyap[1], Muskan Singh[2], Laxmi V[3]

[1]Student, Dept. of Information Science and Engineering, BNM Institute of Technology, Karnataka, India
[2] Student, Dept. of Information Science and Engineering, BNM Institute of Technology, Karnataka, India
[3] Assistant Professor, Dept. of Information Science and Engineering, BNM Institute of Technology, Karnataka, India

---***---

Abstract – *Cryptocurrencies are digital money in which, as opposed to a centralized authority, a decentralized system use encryption to verify transactions and keep records. Decentralization is the transfer of power and authority from a central entity—which might be an individual, business, or group—to a decentralized network. High price volatility of crypto currencies creates a huge impact on international trade. There has been significant jump in price of various currencies like Bitcoin, Ethereum, Litecoin, Dogecoin which has attracted a lot of investors. A lot of research has been done using ARIMA model, SVM and different machine learning techniques but none of them could give as promising results as deep learning models i.e., GRU and LSTM implemented in this paper. Enforcing forecasting models based on deep learning is the key goal. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are two deep learning models that are employed in this study to deal with extreme price swings and produce reliable findings. The two models are further contrasted, and the best of the two is used to anticipate prices. Using two separate methodologies for error prediction, mean absolute percentage error (MAPE) and root mean square error (RMSE), it can be seen that GRU outperforms LSTM for the majority of cryptocurrencies.*

Key Words: Recurrent Neural Network (RNN), LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), Deep Learning (DL)

1. INTRODUCTION

For the past several years, predicting the price of cryptocurrencies has been an exciting field of study. As a leader in the blockchain financial renaissance, Bitcoin [1] [2] contributes significantly to the market capitalization of cryptocurrencies. Cryptocurrency pricing forecast can provide assistance to both beginners and experienced investors in making the right investment decisions to maximize profits while also supporting policy decisions and financial analysts to study cryptocurrency market behavior. Cryptocurrency forecasts are considered time series problems, just like stock price forecast. Crypto price prediction has been a challenging task due to its high volatility. Volatility refers to the sudden change in market sentiments. The market sentiment is influenced by various factors such as public relations, political system, market policies. Factors such as international relations can determine economic role of crypto on different market strategies. The current study's objective is to reduce risk for investors and policymakers by better forecasting the price of cryptocurrencies using deep learning models.

2. RELATED WORK

Studies that use machine learning to predict cryptocurrencies are inadequate, especially in the techniques of deep learning. According to a survey done in the year 2016, over 600 articles have been published in the area. This paper discusses about an experiment conducted on more than six cryptocurrencies to estimate their closing prices using various methodologies, as well as the requirements and evaluation of RNN and its system architectural design.

In [3], it applied machine learning techniques like support vector machine (SVM), recurrent neural network (RNN), and artificial neural network (ANN), as well as k-means clustering, to predict the price of bitcoin using a variety of attribute selection techniques to identify the most crucial features. The fact that this research exclusively focuses on investors, however, is a drawback. Policymakers should be viewed as essential participants in the process since cryptocurrencies have the potential to alter the volatility of the global economy.

In [4], the research focused on computational intelligence methods, particularly the hybrid neuro-fuzzy control of predicting bitcoin exchange rates. This model uses the neuro-fuzzy method and the ANN. For the purpose of producing reliable prognosis findings, the project's deep learning (DL), reinforcement learning (RL), and current deep neural network (NN) networks were combined in [5] to provide an extensive study framework for direct financial signal representation and trading. Using data from the stock market and futures markets, the proposed strategy was later validated. [6] claimed that all market evaluations of virtual currencies are directly or indirectly influenced by socially formed beliefs about virtual money on a network like Twitter. This study aimed to determine the link between positive and negative attitudes that were gathered from Twitter in order to anticipate the fluctuating value of bitcoin using sentiment analysis.

Various machine learning techniques have been applied in [7] to better correctly estimate the value of bitcoin. Using a machine learning technique analogous to LSTM, it improved the estimate for future stock prices in [8]. Because it is a crucial stage in forecasting stock prices and other models of financial forecasting outcomes, the study primarily focuses on time series forecasts. Using data from the stock market and commodities futures markets, they validate the suggested methodology. Additionally, when compared to the current ARIMA model, the LSTM algorithm yields effective and precise outcomes.

The direction of the Bitcoin value in USD was predicted in [9] using Bayesian optimized RNN and LSTM. The Autoregressive Integrated Moving Average (ARIMA) model was also included in order to compare the different deep learning approaches.

3.KEY TECHNOLOGIES

3.1 Deep Learning

Deep learning is a AI strategy that teaches computers to learn and do what human beings do normally. DL describes a family of learning algorithms rather than a single method that can be used to learn complex prediction models and for example be used in stock and crypto prediction. Compared to the Conventional Neural Network (CNN) which comprises 2-3 hidden layers, the deep neural networks may sum up to one hundred layers or more.

3.2 Recurrent Neural Network

RNNs are deep neural networks characterized as repetitive associations between the inputs and outputs of neurons or layers that can train sequence designed to capture relevant temporary data and temporal sequence information. When it comes to learning long sequences, it has recently become popular in deep learning because it overcomes the limitations of existing neural network architectures. The two common RNN networks used for forecasting are LSTMs and GRUs, which are described in the next section.

4. PROPOSED METHOD

In order to estimate the closing price of cryptocurrencies on an intraday basis by recognizing and analyzing pertinent factors by the model itself, the proposed technique compares and contrasts two alternative deep learning-based prediction models. Figure 1 depicts the recommended technique's process design. The Kaggle dataset is downloaded first. Max-Min Normalization is used to scale the data after that, and it is then further pre-processed using data augmentation, binning, dimensionality reduction, etc. Further, the dataset is split randomly into 70% training and 30% testing. 30% of the data is kept at training to test the accuracy of the techniques precisely. After implementing both the techniques for price forecasting, determination of which model performs better and selection of appropriate method to obtain better efficacy. Deep learning algorithms like LSTM and GRU, which are the most recent and effective methods for forecasting the closing price of cryptocurrencies, have been proposed in this study. Both models provide predictions about the final price of cryptocurrency.

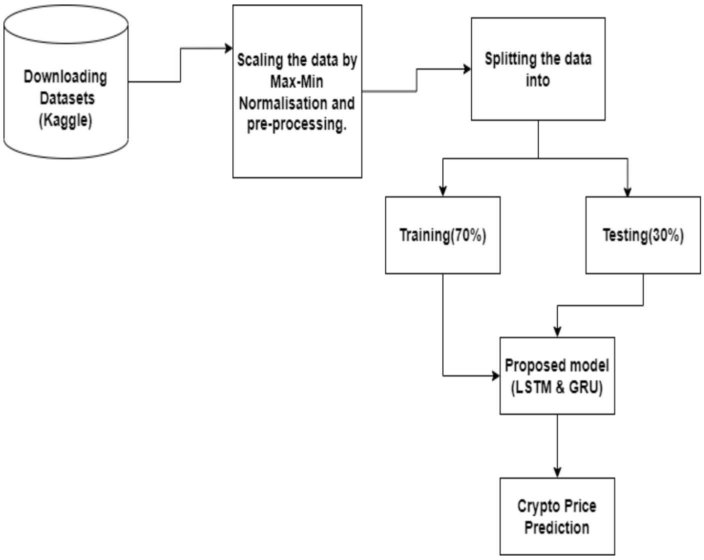

Fig -1: Block Diagram of Proposed Method

4.1 Dataset Preparation

The most common technique for gathering, collecting, categorizing, and organizing information is data preparation, which encompasses data visualization, analysis, and information mining with deep learning applications. To forecast bitcoin values, the proper dataset must be made available. This study used daily bitcoin prices, and the dataset was gathered via the Kaggle website at https://www.kaggle.com. From 2013 until 2021, daily access to pricing history is available. The dataset under examination includes seven factors, including date, opening price, high price, low price, and closing price; volume; and market potential of publicly traded outstanding shares, which are used to anticipate the closing price of crypto currencies.

4.2 LSTM

Long-term dependencies are the problems for which the desired output is dependent on input that are present far in the past. LSTMs are models especially built to solve the long-term dependency problem. It is the default behavior of LSTM to learn information for long interval of time. All recurrent neural networks are in the form of chain of repetitive neural network modules. For standard RNNs, the repeating module has a basic structure called a single layer of "tanh". The deep learning LSTM neural networks resolve the problems of declining gradient descent. The problem is overcome by adding memory cells and a gating mechanism in place of the RNN's nodes.

Figure 2 shows the LSTM block diagram. Before passing on the long-term and short-term information to the following cell, the LSTM cells employ the gates to control the information that is to be maintained or that is to be discarded during loop operation. The three gates serve as filters that exclude irrelevant, picked, and undesirable information.

The gates and layers that make up the LSTM are as follows:

Input Gate: The input gate is utilized to update the state of the cell. The sigmoid function determines which values will be changed based on the output from the previously concealed state and the current input. By changing the values to be between 0 and 1, where 0 denotes that information is not important and 1 denotes that information is significant and must be maintained, the sigmoid function calculates the updated value. The tanh function is then given the hidden state and current input. Tanh function is used to squish the values between -1 and 1 for network regulation. The sigmoid output is multiplied by the additional tanh output. Which information from tanh is significant and ought to be retained is determined by the resulting sigmoid output.

Forget Gate: To determine whether to keep or discard the information, forget gate is employed. The sigmoid function receives input from the previous hidden state and the current input and outputs values between 0 and 1. The values which are approximately 0 are the information to get discarded, and the values which are approximately 1 are the information to keep.

Output Gate: The output gate is used to decide which value should be the next hidden state and which information is to be passed based on previous inputs. It also uses hidden state for prediction. Steps followed in output gate is: The sigmoid function receives two inputs: the most recent hidden state and the most recent input. After that, the newly transformed cell state is added to the tanh function. Additionally, the sigmoid output is repeatedly coupled with the result of the tanh function to determine what information from the hidden state should be carried. The concealed state is the output. Next time step includes the new hidden state as well as the new cell state.

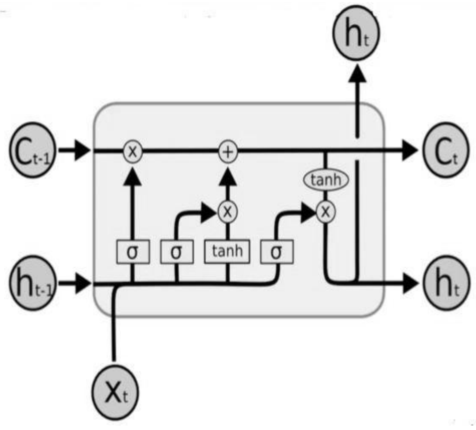

Fig -2: Block Diagram of LSTM

4.3 GRU

A new generation of recurrent neural networks, the GRU depicted in Figure 3, is comparable to the LSTM. The GRU deletes the cell state and sends the data to subsequent neurons using the concealed status. GRU's has fewer tensor operation than LSTM's therefore it is a little faster to train the model according to the dataset. It has a simple design and lesser number of gates compared to LSTM.

The gates and layers that make up the LSTM are as follows:

Update Gate: An LSTM's update gate functions in the same way as its forget and input gates put together. It makes decisions on which data to keep and which to delete.

Reset Gate: Another gate used to control how much prior knowledge is to be forgotten is the reset gate.

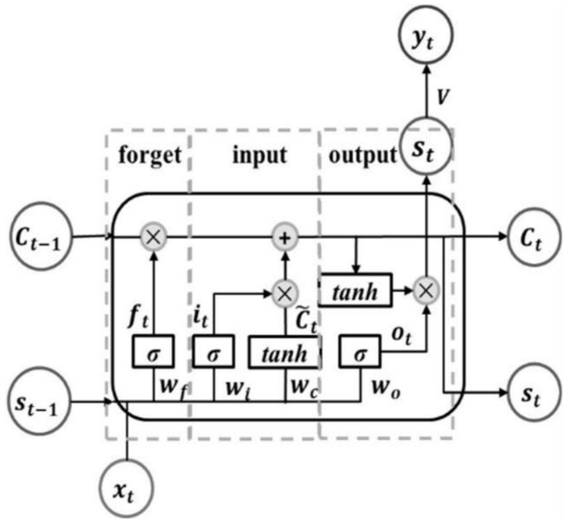

Fig -3: Block Diagram of GRU

4.4 Activation Functions present in LSTM and GRU

Sigmoid: Similar to tanh activations, sigmoid activations are seen in Gates. Instead of squishing numbers between -1 and 1, it squishes values between 0 and 1. Because every integer multiplied by 0 equals 0, values are overlooked or "forgotten" when using this procedure to update or forget data. Any integer multiplied by one produces the same value as a result; consequently, the value is "preserved." The network can learn which information in a dataset is vital and should be retained and which information is unimportant and should be deleted.

Tanh: Tanh activation aids in controlling the values that are transmitted across the network. Values are always squeezed to fall between -1 and 1 by the tanh function. Due of the multiple arithmetic operations created, vectors travelling through a neural network must go through several modifications. [9]

5. RESULTS AND ANALYSIS

The two suggested deep learning models, the LSTM and GRU, are trained. After the models have been trained, the results are examined using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) to identify which model has the highest accuracy. It is noted in the tests that the LSTM model takes longer duration to compile than the GRU model. From the acquired values and the graphs plotted it is evident that GRU model converges faster than LSTM model and is steadier.

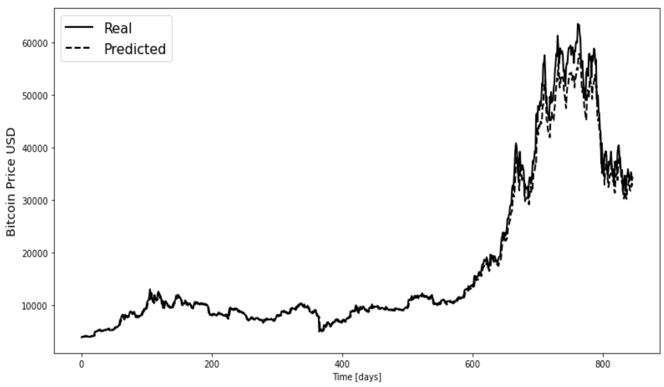

Fig 4.a.: Closing Prices of Bitcoin predicted by LSTM

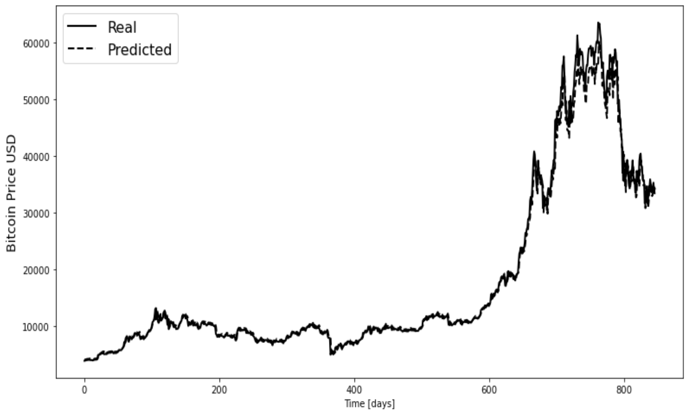

Fig 4.b.: Closing Prices of Bitcoin predicted by GRU

Figure 4.a. represents the closing price of Bitcoin as forecasted by the LSTM model and Figure 4.b. represents the closing price of Bitcoin as forecasted by GRU. From both the figures, it can be understood that GRU has more efficacy in forecasting the crypto price as compared to LSTM.

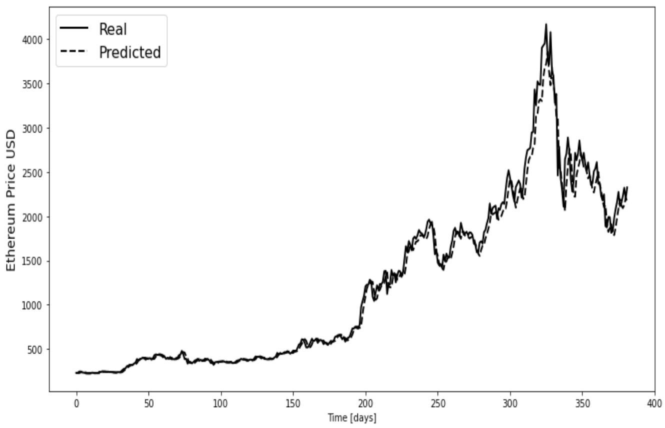

Fig 5.a.: Closing Prices of Ethereum predicted by LSTM

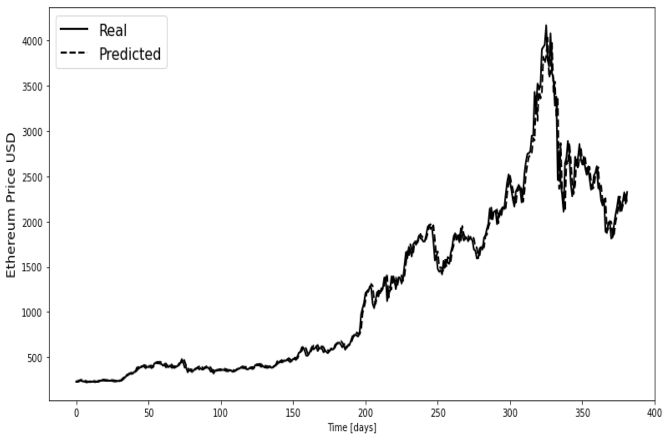

Fig 5.b.: Closing Prices of Ethereum predicted by GRU

Figure 5.a. represents the closing price of Ethereum as forecasted by the LSTM model and Figure 5.b. represents the closing price of Ethereum as forecasted by GRU. From both the figures, it can be understood that GRU has more efficacy in forecasting the crypto price as compared to LSTM.

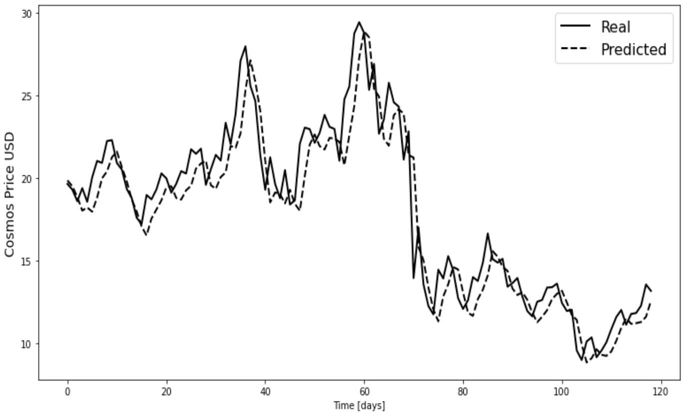

Fig 6.a.: Closing Prices of Cosmos predicted by LSTM

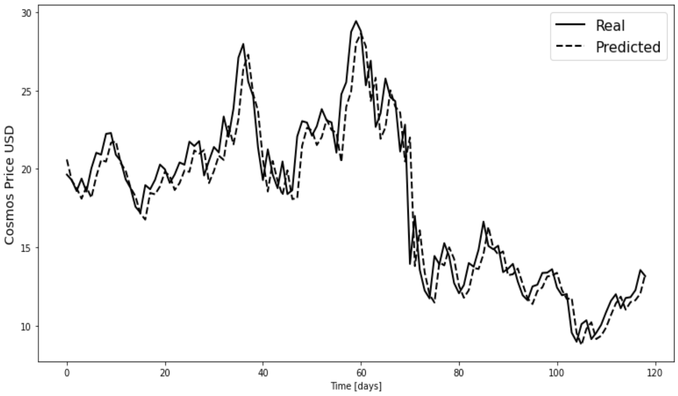

Fig 6.b.: Closing Prices of Cosmos predicted by GRU

Figure 6.a. represents the closing price of Cosmos as forecasted by the LSTM model and Figure 6.b. represents the closing price of Cosmos as forecasted by GRU. From both the figures, it can be understood that GRU has more efficacy in forecasting the crypto price as compared to LSTM.

5.1 Performance Measures

In this paper, GRU-based forecasting model is more appropriate than LSTM for most of the cryptocurrencies to predict time series data of highest price instability. Two error finding techniques have been used:

Root Mean Square Error: It is the standard deviation of the residuals (prediction errors). Residuals are basically a measure of how far the data points lie from the line of regression. It is a measure of how spread out these residuals are which tells how concentrated the data is around the line of best fit. It is commonly used in climatology, forecasting, and regression analysis to verify experimental results.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

Mean Absolute Percentage Error: It measures accuracy of a forecasting system. It measures this accuracy in the form of percentage, and can be calculated as the average absolute percent error for each time period minus actual values whole divided by actual values.

$$MAPE = \sum_{t=1}^{N} \left| \frac{observed_t - predicted_t}{observed_t} \right| \times \frac{100}{N}$$

Table -1: Error Rate

Type of Cryptocurrency	LSTM		GRU	
	RMSE	MAPE	RMSE	MAPE
Bitcoin	0.048	0.029	0.046	0.026
Ethereum	0.030	0.017	0.028	0.016
Dogecoin	0.035	0.011	0.034	0.011
Binance Coin	0.044	0.025	0.041	0.024
LiteCoin	0.026	0.014	0.032	0.018
Cosmos Coin	0.060	0.045	0.060	0.044
Cardano Coin	0.032	0.049	0.058	0.040

Table 1 shows the RMSE and MAPE values for various cryptocurrencies. From table 1, it is evident that GRU has lesser error rate when compared to LSTM model.

6. CONCLUSION AND FUTURE SCOPE

Crypto is decentralized method of virtual money. It assumes a crucial part in unrestricted economy. One of the major advantages of crypto currency is that it eradicates third party intermediary among the users. The main purpose behind this work is to predict prices of various crypto currencies present in the market, efficiently. There has been numerous research on different methods for crypto price prediction. There are many factors such as social, economic, geographical, political which effects crypto prices and makes it volatile in nature. Due to high volatility of the time series data, accuracy of crypto prediction is not good. Through the study of various research paper, it is seen that RNN based model provides best accuracy amidst the volatility and all fluctuations. To get accurate results, deep learning models are used which helps in minimizing the risk and makes predictions more stable. The proposed paper compares two

RNN based models i.e., LSTM and GRU. From the experiments conducted, it is observed that GRU model performs better for time series predictions and takes less time to compile. Both the models are known for their recognition in long-term dependencies.

The future scope of the proposed work is to enhance the accuracy of predictions by considering more parameters like public relations, political activity, market policy, etc. into account. Incorporation of Fuzzification is also one of the major future scopes.

REFERENCES

[1] D.L.K. Chuen, Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data, Academic Press, 2015.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008

[3] Dennys CA, Mallqui RAF (2018) Predicting the direction, maximum, minimum and closing prices of daily bitcoin exchange rate using machine learning techniques. Int J Soft Comput (IJSC) 596–606

[4] Atsalakis GS, Atsalaki IG, Pasiouras, F, Zopounidis C (2019) Bitcoin price forecasting with neuro-fuzzy techniques. Elsevier, pp 770–780

[5] Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press

[6] Pant DR, Neupane P, Poudel A, Pokhrel AK, Lama BK (2018) Recurrent neural network based bitcoin price prediction by twitter sentiment analysis. IEEE, pp 128–132

[7] Phaladisailoed T, Numnonda T (2018) Machine learning models comparison for bitcoin price prediction. In: 2018 10th international conference on information technology and electrical engineering (ICITEE). IEEE, pp 506–511

[8] Nivethitha P, Raharitha P (2019) Future stock price prediction using LSTM machine learning algorithm. Int Res J Eng Technol (IRJET) 1182–1186

[9] https://mlcheatsheet.readthedocs.io/en/latest/activation_functions.html#tanh

[10] McNally S, Roche J, Caton S (2018) Cryptocurrency forecasting with deep learning chaotic neural networks. IEEE, pp 339–343

[11] Madan I, Saluja S, Zhao A(2015) Automated bitcoin trading via machine learning algorithms. 1–5. http://cs229.stanford.edu/proj2014/Isaac\%

[12] Lahmiri S, Bekiros S (2019) Cryptocurrency forecasting with deep learning chaotic neural networks. Elsevier, 35–40

[13] Saxena A, Sukumar TR (2018) Predicting bitcoin Price using lstm and compare its predictability with Arima model. Int J Pure Appl Math 2591–2600