

# Stock Prediction using Hybrid ARIMA and GRU Models

Rahul Mangalampalli

Department of Information Technology Engineering  
D. Y. Patil College of Engineering, Ambi  
Pune, India

Pawankumar Khetre

Department of Information Technology Engineering  
D. Y. Patil College of Engineering, Ambi  
Pune, India

Vivek Pandey

Department of Information Technology Engineering  
D. Y. Patil College of Engineering, Ambi  
Pune, India

Vaibhav Malviya

Department of Information Technology Engineering  
D. Y. Patil College of Engineering, Ambi  
Pune, India

**Abstract**— Prediction of stocks requires a lot of knowledge on market share values and trends. This knowledge can be obtained by experience in this particular field. For a normal human it requires a lot of time and energy to gain experience to predict trends in stock prices. With advancement in technology, machine learning algorithms keep the capability of predicting trends in stocks because of the huge computational capacity which is available nowadays. In this paper, hybrid ARIMA (Auto Regressive-Integrated-Moving Average)-GRU(Gated-Recurrent-Unit) model has been proposed which learns continuously from the history of stocks invested, sold and brought by the clients. ARIMA-GRU model identifies patterns, analyzes which kind of data is suitable for prediction and then predicts suitable value. A detailed information about working of this model, parameters it considers and type of values it predicts have been discussed in this paper.

**Keywords**— Autoregressive integrated moving average, gated recurrent unit, Stock price prediction, Hybrid models

## I. INTRODUCTION

Stock Prediction plays a vital role in finance and economics. Today many investors and organizations invest their money in stocks but they don't have any idea about the future results. Today many people invest billions of currencies on the stocks expecting the profit after every single stock purchased. Depending on the behavior of market there are ups and downs in the profit. Sometimes there is a great amount of profit and sometimes there is a loss. So, to overcome this issue we are introducing a concept which is Hybrid Models of Machine Learning and Deep Learning Models for Stock Prediction. There are many charges for investing in the stock market the main is the brokerage for the brokers who help the investors to invest in the stock market. Our idea is to implement it in such a way that there should be no brokerage or other charges included for an individual to invest in stock market. This concept would help those investors and organizations to invest their money in a right place as well as they can have a full glance of their investments and future scope of their investments. Many effective algorithms have been introduced to make efficient predictions but most of them failed after very short period due to growing uncertainties in the share market. Uncertainties arise due to

the development of many industries and likeliness of common people investing into their interested fields. Hence it becomes very difficult for any algorithm to decide constant parameters to judge those stock prices. Hybrid ARIMA (Auto Regressive-Integrated-Moving Average)-GRU(Gated-Recurrent-Unit) model has been proposed which learns continuously from the history of stocks invested, sold and brought by the clients. ARIMA-GRU model identifies patterns, analyzes which kind of data is suitable for prediction and then predicts suitable value. A detailed information about working of this model, parameters it considers and values it predicts have been discussed in this paper.

## II. LITERATURE SURVEY

Using Neural Networks to Forecast Stock Market Prices, Ramon Lawrence.

This paper is a survey on the application of neural networks in forecasting stock market prices. With their ability to discover patterns in nonlinear and chaotic systems, neural networks offer the ability to predict market directions more accurately than current techniques. Common market analysis techniques such as technical analysis, fundamental analysis, and regression are discussed and compared with neural network performance. Also, the Efficient Market Hypothesis (EMH) is presented and contrasted with chaos theory and neural networks. Finally, future directions for applying neural networks to the financial markets are discussed [1].

Stock Market Prediction Using Hybrid Approach, Vivek Rajput, Sarika Bobde.

The objective of this paper is to construct a model to predict stock value movement using the opinion mining and clustering method to predict National Stock Exchange (NSE). It used domain specific approach to predict the stocks from each domain and taken some stock with maximum capitalization. Topics and related opinion of shareholders are automatically extracted from the writings in a message board by utilizing our proposed strategy alongside isolating clusters of comparable sort of stocks from others using clustering algorithms. Proposed methodology will give two output set i.e. one from sentiment analysis and another from clustering-based prediction with respect to some specialized parameters

of stock exchange. By examining both the results an efficient prediction is produced. In this paper stocks with maximum capitalization within all the important sectors are taken into consideration for empirical analysis [2].

Hybrid ARIMA-BPNN Model for Time Series Prediction of the Chinese Stock Market, Li Xiong, Yue Lu.

Stock price prediction is a challenging task owing to the complexity patterns behind time series. Autoregressive integrated moving average (ARIMA) model and back propagation neural network (BPNN) model are popular linear and nonlinear models for time series forecasting respectively. The integration of two models can effectively capture the linear and nonlinear patterns hidden in a time series and improve forecast accuracy. In this paper, a new hybrid ARIMA-BPNN model containing technical indicators is proposed to forecast four individual stocks consisting of both main board market and growth enterprise market in software and information services sector [3].

Deep Learning for Stock Market Prediction Using Technical Indicators and Financial News Articles, Manuel R. Vargas, Carlos E. M. dos Anjos, Gustavo L G. Bichara, Alexandre G. Evsukoff.

This work uses deep learning models for daily directional movements prediction of a stock price using financial news titles and technical indicators as input. A comparison is made between two different sets of technical indicators, set 1: Stochastic (%K), Stochastic (%D), Momentum, Rate of change, William's (%R), Accumulation/Distribution (A/D) oscillator and Disparity 5; set 2: Exponential Moving Average, Moving Average Convergence-Divergence, Relative Strength Index, On Balance Volume and Bollinger Bands. Deep learning methods can detect and analyze complex patterns and interactions in the data allowing a more precise trading process. Experiments has shown that Convolutional Neural Network (CNN) can be better than Recurrent Neural Networks (RNN) on catching semantic from texts and RNN is better on catching the context information and modeling complex temporal characteristics for stock market forecasting. So, there are two models compared in this paper: a hybrid model composed by a CNN for the financial news and a Long Short-Term Memory (LSTM) for technical indicators, named as SI-RCNN; and a LSTM network only for technical indicators, named as I-RNN. The output of each model is used as input for a trading agent that buys stocks on the current day and sells the next day when the model predicts that the price is going up, otherwise the agent sells stocks on the current day and buys the next day. The proposed method shows a major role of financial news in stabilizing the results and almost no improvement when comparing different sets of technical indicators [4].

Financial Indices Modelling and Trading utilizing deep learning techniques, Marios Mourelatos, Thomas Amorgianiotis, Christos Alexakos, Spiridon Likothanassis.

Prediction and modelling of the financial indices is a very challenging and demanding problem because its dynamic, noisy and multivariate nature. Modern approaches have also to challenge the fact that they are dependencies

between different global financial indices. All this complexity in combination with the large volume of historic financial data raised the need for advanced machine learning solutions to the problem. This article proposes a Deep Learning approach utilizing Long Short-Term Memory (LSTM) Networks for the modelling and trading of financial indices [5].

Hybrid Deep Learning Models for Stock Prediction, Mohammad Asiful Hossain, Rezaul Karim, Ruppa Thulasiram, Neil D B. Bruce, Yang Wang.

Stock market prediction has always caught the attention of many analysts and researchers. Popular theories suggest that stock markets are essentially a random walk and it is a fool's game to try and predict them. Predicting stock prices is a challenging problem in itself because of the number of variables which are involved. This paper reviews all these points [6].

Stock index forecasting based on a hybrid model, J.J. Wang, J. Z. Wang, Z. G. Zhang, and S. P Guo.

This paper examines the prediction performance of ARIMA and artificial neural networks model with obtained stock information from New York Stock Exchange. The empirical results obtained reveal the prevalence of neural networks model over ARIMA model. The findings further resolve and clarify contradictory opinions reported in literature over the prevalence of neural networks and ARIMA model and the other way around [7].

A moving-average filter-based hybrid ARIMA-ANN model for forecasting time series data, C. Narendra Babu and B. Eswara Reddy.

A suitable combination of linear and nonlinear models provides a lot of correct prediction model than a individual linear or nonlinear model for foretelling statistic knowledge originating from numerous applications. The linear autoregressive integrated moving average (ARIMA) and nonlinear artificial neural network (ANN) models s explored during this paper to plan a brand new hybrid ARIMA-ANN model for the prediction of your time series knowledge [8].

Supervised Sequence Labelling with Recurrent, Graves.

This paper provides the background material and literature review for supervised sequence labelling. Brief reviews are done on supervised learning in general and covers the classical, non-sequential framework of supervised pattern classification. It also defines supervised sequence labelling, and describes the different classes of sequence labelling task that arise under different assumptions about the label sequences [9].

Backpropagation through time: what it does and how to do it, P.J Werbos.

This paper reviews the basic idea of backpropagation, a simple method which is being widely used in areas like pattern recognition and fault diagnosis. It further expands the idea of dealing with recurrent networks,

systems involving simultaneous equations. The chain rule for ordered derivatives, the theorem which underlies backpropagation is briefly discussed [10].

A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices, Rudra Kalyan Nayaka, Debahuti Mishraa, Amiya Kumar Rathb.

This paper proposes a hybridized framework of Support Vector Machine (SVM) with K-Nearest Neighbor approach for Indian stock market indices prediction. The objective of this paper is to get in-depth knowledge in the stock market in Indian Scenario with the two indices such as, Bombay Stock Exchange (BSE Sensex) and CNX Nifty using technical analysis methods and tools such as predicting closing price, volatility and momentum of the stock market for the available data. This hybrid model uses SVM with different kernel functions to predict profit or loss, and the output of SVM helps to compute best nearest neighbor from the training set to predict future of stock value in the horizon of 1 day, 1 week and 1 month. The proposed SVM and KNN based prediction model is experienced with the above mentioned distinguished stock market indices and the performance of proposed model has been computed using Mean Squared Error and also been compared with recent developed models such as FLIT2NS and CEFLANN respectively [11].

A Hybrid Fuzzy Time Series Model Based on ANFIS and Integrated Nonlinear Feature Selection Method for Forecasting Stock, Chung-Ho Su, Ching-Hsue Cheng.

Forecasting stock price is a hot issue for stock investors, dealers and brokers. However, it's difficult to find out the best time point to buy or to sell stock, due to many variables will affect the stock market, and stock dataset is time series data. Therefore, many time series models have been proposed for forecasting stock price, furthermore the previous time series methods still have some problems. Hence, this paper proposes a novel ANFIS (Adaptive Neuro Fuzzy Inference System) time series model based on integrated nonlinear feature selection (INFS) method for stock forecasting [12].

Hybrid nonlinear adaptive scheme for stock market prediction using feedback FLANN and factor analysis, C.M. Anish, Babita Majhi.

Accurate and effective stock price prediction is very important for potential investors in deciding investment strategy. Data mining techniques have been applied to stock market prediction in recent literature. Factor analysis (FA), a powerful statistical attributes reduction technique, is chosen to select the inputs of the model from the raw data. A feedback type of the functional link artificial neural network (FFLANN) with recursive least square (RLS) training is proposed as a potential prediction model [13].

A Hybrid Time Series Model based on AR-EMD and Volatility for Medical Data Forecasting - A Case Study in the Emergency Department, Liang-Ying Wei Deng-Yang Huang Shun-Chuan Ho Jyh-Shyan Lin Hao-En Chueh Chin-Sung Liu Tien-Hwa Ho.

Time series methods have been applied to forecast clinical data, such as daily patient number forecasting for emergency medical centers. However, the application of conventional time series models needs to meet the statistical assumptions, and not all models can be applied in all datasets. Most of the traditional time series models use a single variable for forecasting, but there are many noises involutedly in raw data that are caused by changes in weather conditions and environments for daily patient number forecasting [14].

Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction, Ayodele Ariyo Adebisi, Aderemi Oluyinka Adewumi and Charles Korede Ayo.

This paper examines the forecasting performance of ARIMA and artificial neural networks model with published stock data obtained from New York Stock Exchange. The empirical results obtained reveal the superiority of neural networks model over ARIMA model. The findings further resolve and clarify contradictory opinions reported in literature over the superiority of neural networks and ARIMA model and vice versa [15].

On the properties of neural machine translation: Encoder-decoder approaches, K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio.

Neural machine translation is a relatively new approach to statistical machine translation based purely on neural networks. The neural machine translation models often consist of an encoder and a decoder. The encoder extracts a fixed-length representation from a variable-length input sentence, and the decoder generates a correct translation from this representation. In this paper, this paper focus on analyzing the properties of the neural machine translation using two models; RNN Encoder-Decoder and a newly proposed gated recursive convolutional neural network. It shows that the neural machine translation performs relatively well on short sentences without unknown words, but its performance degrades rapidly as the length of the sentence and the number of unknown words increase. Furthermore, this paper finds that the proposed gated recursive convolutional network learns a grammatical structure of a sentence automatically [16].

Time series forecasting using a hybrid ARIMA and neural network model, G. P. Zhang.

Autoregressive integrated moving average (ARIMA) is one of the popular linear models in time series forecasting during the past three decades. Recent research activities in forecasting with artificial neural networks (ANNs) suggest that ANNs can be a promising alternative to the traditional linear methods. ARIMA models and ANNs are often compared with mixed conclusions in terms of the superiority in forecasting performance. In this paper, a hybrid methodology that combines both ARIMA and ANN models is proposed to take advantage of the unique strength of ARIMA and ANN models in linear and nonlinear modeling. Experimental results with real data sets indicate that the combined model can be an effective way to improve

forecasting accuracy achieved by either of the models used separately. [17].

Predicting Prices of Stock Market using Gated Recurrent Units (GRUs) Neural Networks, Mohammad Obaidur Rahman, Md. Sabir Hossain, Ta-Seen Junaid, Md. Shafiu Alam Forhad, Muhammad Kamal Hossen.

In this paper, proposed a model is designed to predict the future prices of the stock market using Gated Recurrent Units (GRUs) neural networks. The paper depicts changed internal structure of GRUs in order to remove local minima problem, reduce time complexity and other problems of stochastic gradient descent as well as improve the efficiency. It has used minibatch gradient descent, is a good trade-off between stochastic gradient descent and batch gradient descent. Then evaluated result by calculating the root mean square error on the various dataset. After extensive experiments on the real-time dataset, proposed method predicted the future prices successfully with good accuracy. [18].

Learning long-term dependences with gradient descent is difficult", Bengio, Yoshua, S. Patrice, F. Paolo.

Recurrent neural networks can be used to map input sequences to output sequences, such as for recognition, production or prediction problems. However, practical difficulties have been reported in training recurrent neural networks to perform tasks in which the temporal contingencies present in the input/output sequences span long intervals. We show why gradient based learning algorithms face an increasingly difficult problem as the duration of the dependencies to be captured increases. These results expose a trade-off between efficient learning by gradient descent and latching on information for long periods. Based on an understanding of this problem, alternatives to standard gradient descent are considered [19].

Applied attention-based LSTM neural networks in stock prediction., Cheng, Li-Chen, Yu-Hsiang Huang, and Mu-En Wu.

Prediction of stocks is complex due to dynamic, complex, and chaotic environment of the stock market. several studies predict that stock value movements are using deep learning models. though the main mechanism has gained quality recently in neural computational translation, little focus has been dedicated to attention-based deep learning models for stock prediction [20].

Short term stock price prediction using deep learning, Khare, Kaustubh.

Short - term price movements, contribute a substantial live to the unpredictability of the securities exchanges. Accurately predicting the price fluctuations available market may be a huge economical advantage. The aforesaid task is mostly achieved by analyzing the corporate, this can be known as fundamental analysis. Another technique, that is undergoing tons of analysis work recently, is to form a predictive algorithmic model using machine learning [21].

Artificial Neural Networks architectures for stock price prediction: comparisons and applications, L. Di Persio and O. Honchar.

Artificial Neural Network (ANN) approach to predict stock market indices, particularly with respect to the forecast of their trend movements up or down. Exploiting different Neural Networks architectures, this paper provides numerical analysis of concrete financial time series. In particular, after a brief resume of the existing literature on the subject, it considers the Multi-layer Perceptron (MLP), the Convolutional Neural Networks (CNN), and the Long Short-Term Memory (LSTM) recurrent neural networks technique [22].

Classification-based Financial Markets Prediction using Deep Neural Networks, Matthew Dixon, Diego Klabjan, Jin Hoon Bang.

Deep neural networks (DNNs) are powerful types of artificial neural networks (ANNs) that use several hidden layers. They have recently gained considerable attention in the speech transcription and image recognition community for their superior predictive properties including robustness to overfitting. However, their application to algorithmic trading has not been previously researched, partly because of their computational complexity. This paper describes the application of DNNs to predicting financial market movement directions. In particular, we describe the configuration and training approach and then demonstrate their application to back testing a simple trading strategy over 43 different Commodity and FX future mid-prices at 5-minute intervals. All results in this paper are generated using a C++ implementation on the Intel Xeon Phi co-processor which is 11.4x faster than the serial version and a Python strategy back testing environment both of which are available as open source code written by the authors [23].

Dynamic Business Network Analysis for Correlated Stock Price Movement Prediction, Wenping Zhang, Chunping Li, Yunming Ye, Wenjie Li and Eric W.T. Ngai.

This paper discusses about a novel business network-based model can help predict directional stock price movements by considering both influential business relationships and Twitter sentiment [24].

Optimizing Stock Market Price Prediction using a Hybrid Approach Based on HP Filter and Support Vector Regression, Meryem Ouahilal, Mohammed El Mohajir, Mohamed Chahhou, Badr Eddine El Mohajir.

Predicting stock prices is an important task of financial time series forecasting, which is of great interest to stock investors, stock traders and applied researchers. Many machine learning techniques have been used in recent times to predict the stock price, including regression algorithms which can be useful tools to provide good accuracy of financial time series forecasting. In this paper, a novel hybrid approach which combines Support Vector Regression and Hodrick-Prescott filter in order to optimize the prediction of stock price has been proposed [25].

### III. PROPOSED METHOD

#### 3.1 Auto-Regressive Moving Average

It is made up of three different models, the three different parts are mainly AR(Auto Regression), MA(Moving Average) and the degree of ordinary differencing and the combined equation for the above three parts will be [7]:

$$(1 - B)Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + \dots + \varphi_p Y_{t-p} + \varepsilon_t - \varepsilon_{t-1}\theta_1 - \varepsilon_{t-2}\theta_2 - \varepsilon_{t-3}\theta_3 + \dots + \varepsilon_{t-q}\theta_q \quad (1)$$

Where

- $B =$  Backshift Operator
- $Y_t =$  Value at time  $t$
- $\delta =$  Bias
- $\varphi =$  Auto -Regressive parameter
- $\varepsilon_t =$  Random error at  $t$
- $\theta =$  Moving Average Parameters

The backshift operator  $B$  is defined to perform shifting of time series data ( $Y$ ) by one period.

$$Y_t B = Y_{t-1} \quad (2)$$

Multiplication with higher degree of  $B$ , gives a backward shift value more than 1 period.

$$Y_t B^2 = Y_{t-2} \quad (3)$$

$$Y_t B^n = Y_{t-n} \quad (4)$$

The first-difference of the series has been shown according to the shift operator  $B$ . Let us assume  $y$  is the first difference of  $Y$ . So, at time  $t$ ,

$$y_t = Y_t - Y_{t-1} = Y_t - B Y_t = (1 - B)Y_t \quad (5)$$

The original series  $Y$  is multiplied with the factor of  $1-B$  to obtain differenced series  $y$ . Now, if assume that  $z$  is the first difference of  $y$ , which makes  $z$  as the second difference of  $Y$ , which gives,

$$z_t = y_t - y_{t-1} = (1 - B)y_t = (1 - B)((1 - B)Y_t) = (1 - B)^2 Y_t \quad (6)$$

By multiplying factor of  $(1 - B)^2$ , the second difference of  $Y$  is obtained. In general, the  $d^{th}$  difference of  $Y$  would be obtained by multiplying by a factor of  $(1 - B)^d$ .

The ARIMA modeling procedures are determined through the Box- Jenkins model building methodology,

- (a) identifying the degree of differencing to transform the time series data into stationary.
- (b) estimating the model parameters by auto correlation function (ACF) and partial ACF (PACF).
- (c) checking the degree of fitting on R square maximum principle and Bayesian Information Criterion (BIC) minimum principle then achieved predicted data and noise residuals [8].

#### 3.2 Gated Recurrent Unit

[9]Engaging technique to resolve in machine learning tasks have recently shown by recurrent neural networks (RNNs). RNN is continuation of a conventional neural network, which may handle a variable-length sequence input. In formally, given input layer of sequence,

$$x = \{x_1, x_2, \dots, x_t\}$$

$h_t$  is hidden layer, output layer,

$$y = \{y_1, y_2, \dots, y_t\}$$

$$h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (7)$$

Where,

$\sigma =$  Activation function (Sigmoid function or Hyperbolic Tangent function)

$W =$  Weight metric

$b =$  bias vector.

The output layer is computed as below formula,

$$y_t = W_{hy}h_t + b_y \quad (8)$$

To train RNN[10], Back Propagation Through Time (BPTT) algorithmic rule is employed. However, it becomes troublesome to coach typical RNNs to capture long-run changes as a result of the vanishing gradient drawback[19]. Therefore, Gated Recurrent Unit (GRU) is employed[16]. Gated Recurrent Unit Recurrent Neural Network (GRU-RNN) address the vanishing problem by replacing hidden node in traditional RNN by GRU node. Every GRU node consists of 2 gates, update gate  $z_t$  and reset gate  $r_t$ . Update gate decides up to what quantity the unit updates its activation, or content. It is computed in equation (9). Reset gate permits to forget the previously computed state, is calculated by equation (10). The hidden layer is computed by equation (12) using  $H_t$  which is calculated by equation (11). In the GRU-RNN we use model parameters including,

$x_t =$  input at time

weight matrices =  $W_z, W_r, W_h, U_z, U_r, U_h$

The following given equations below calculate their respective gates,

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (9)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (10)$$

$$H_t = \tanh(W_h r_t x_t + U_h r_t h_{t-1}) \quad (11)$$

$$h_t = (1 - z_t)h_{t-1} + z_t H_t \quad (12)$$

#### 3.3 Proposed Hybrid of ARIMA-GRU

The hybrid ARIMA-ANN model was pioneered by Zhang [17] in 2003, who assumed time series data are the total of linear and nonlinear parts listed in (13).

$$y_t = L_t N_t \quad (13)$$

where  $L_t$  denotes the linear element and  $N_t$  denotes the nonlinear element. The numerical results displayed higher

prediction accuracy than both ARIMA and ANN models. In 2011, Khashei and Bijari[26] thought the connection between linear and nonlinear parts were not essentially additive and proposed another hybrid ARIMA ANN model, that supported the idea on the assumption that a time series is function of a linear and a nonlinear element as in (14).

$$y_t = f(L_t, N_t) \tag{14}$$

In this model, the forecast values and residuals of ARIMA are treated as inputs of the ANN and through empirical analysis with three real-world time series, this model outperformed Zhang’s model. Based on Khashei and Bijari’s model, this paper established a structure of the proposed ARIMA-GRU model. In the initial stage, the date and opening price are viewed as inputs of ARIMA model and the expected(predicted) opening price are outputs of the model. In the second stage, the outputs from ARIMA model are collected as input variables of GRU model. The output is the trend for next n number of days.

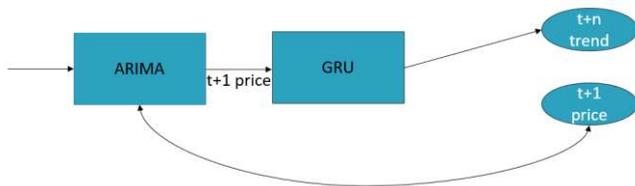


Figure 1: The proposed ARIMA-GRU model

### 3.4 Online Learning

In online learning, training of the system is done by incrementally feeding it data instances sequentially, either individually or by small groups called batches (batch – processing). Each learning step is fast and cheap, so the system can learn about new data continuously. As stock market is a place where new data enters constantly, as it arrives, if the computing resources; once an online learning system has learned about new data instances, it does need them anymore, as the learning process is done, the data instance can be discarded. This can save a huge amount of space.

This method can be used to train systems on huge time series data to that can fit in one machine’s memory. The algorithm loads part of the data, it trains the data batch by batch until the model trains on whole data.

A big challenge with online learning is that if we feed non-representative data to the model, then it’s performance will gradually decline. To avoid this problem, the system needs to be closely monitored and promptly switch off continuous learning whenever a drop in performance is observed. If any anomaly in data is present then that must be avoided.

## IV. INPUTS AND EXPERIMENT RESULTS

The main aim of this paper is to predict a sensible value of a stock belonging to a company or an organization and the trend in its sequence. Hence the inputs given to the model must cleaned thoroughly and preprocessing techniques like

standardization and normalization must be applied before passing them to ARIMA model. Since the given data would be time series, stationarity of the data must be checked by using Augmented Dickey Fuller test.

We have used Google Stock prices data for training our model. The dataset contains date, previous day closing price, current day opening price, current day closing price, current day lowest day and highest prices, change in current day and previous day price, number of deals, volumes of stocks traded and current day collection prices as feature columns.

For our ease of experimentation and to get proper insights of the model we have used only open prices as the feature. The model was trained on total of 1423 data points.

ARIMA is trained to give output for 1 day.i.e. if we give input for price at day t, we’ll get output for t+1 day.

GRU is trained to show trend for 60 days.i.e if we give input for the price at t+1, we’ll get output for next t+1+60 days.

### 4.1 Input Variables And Their Description:

#### 4.1.1 Training Mode:

Variables	Indicators	Description	Shape
ARIMA model inputs	Opening price data	The opening price of day t	(1423,1)
ARIMA model outputs	$\varphi, \delta, \epsilon_t, \theta$	AR, Bias, Residual, MA parameters	(1,1,1,1)
GRU model inputs	Opening price data	The opening price of day t	(1241,60,1)
ARIMA-GRU model outputs	t+1 predicted opening price and trend	The exact predicted value and the trend for next n days.	Trend

#### 4.1.2 Testing Mode:

Variables	Indicators	Description
ARIMA model inputs	Price at t	The opening price of day t
ARIMA model outputs	Price at t+1	The opening price predicted on t+1 day
GRU model inputs	Output of ARIMA at t+1	The opening price predicted on t+1 day
ARIMA-GRU model	t+60 trend	The exact predicted value and the trend for next 60 days.

### 4.2 Results:

Gated Recurrent Units alone can perform an extremely wonderful without any support[27].Hence the hybrid of both ARIMA and GRU can out perform any other RNN-ML algorithm present in the market.This takes care that the given time series data does not overfit the model.Space complexity won’t be a problem at all,because the model would using batch processing for the training purpose of the model.

We have tested on traditional ARIMA-BPNN model and ARIMA-GRU model.

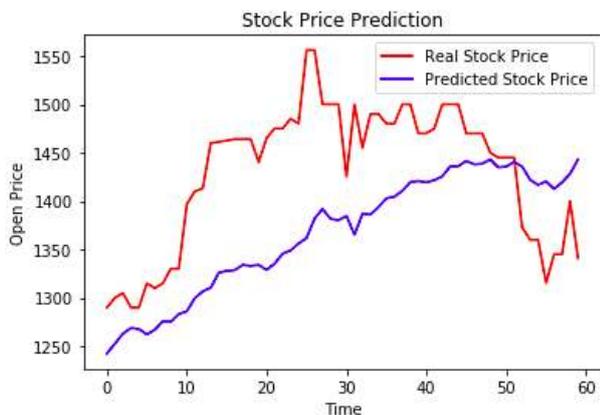


Figure 2: ARIMA-BPNN Testing Result

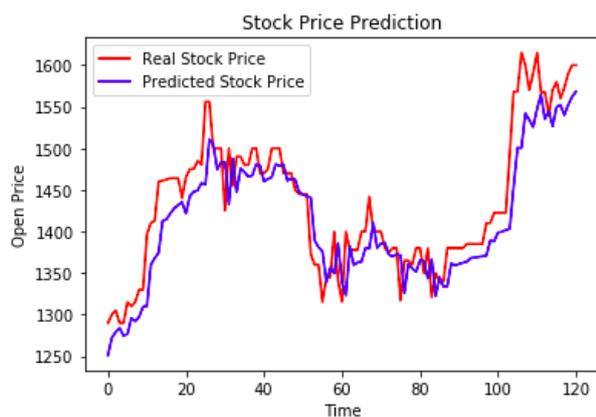


FIGURE 2: ARIMA-GRU TESTING RESULT

## V. CONCLUSION

The proposed model works efficiently when there is no constraint of time complexity, the model needs learn on very huge amount of datasets and come to certain conclusions on how stocks change during the course of time. Once the model is prepared fully with considering each and every anomaly, then it needs to be deployed into the system for further learning and predictions. These model works very perfectly once gets mature and the maturity in its predictions can be observed as it proceeds further in time. As it is a computer program which can learn faster than human, it can get mature very fast.

## REFERENCES

- [1] Ramon Lawrence. "Using Neural Networks to Forecast Stock Market Prices". Neural Networks in the Capital Markets, chapter 10, pages 149–162. John Wiley and Sons, 1995.
- [2] Vivek Rajput, Sarika Bobade. "Stock Market Prediction Using Hybrid Approach". International Conference on Computing, Communication and Automation (ICCCA2016).
- [3] Li Xiong, Yeu Lu (2017). "Hybrid ARIMA-BPNN Model for Time Series Prediction of the Chinese Stock Market". 2017 3rd International Conference on Information Management.
- [4] Manuel R. Vargas, Carlos E.M. dos Anjos, Gustavo L.G. Bichara, Alexandre G. Evsukoff (2018). "Deep Learning for Stock Market Prediction Using Technical Indicators and Financial News Articles". 2018 International Joint Conference on Neural Networks (IJCNN).
- [5] Marios Mourelatos, Thomas Amorgianiotis, Christos Alexakos, Spiridon Likothanassis (2018). "Financial Indices Modelling and Trading Utilizing Deep Learning Techniques". 2018 Innovations in Intelligent Systems and Applications (INISTA).
- [6] Mohammed Asiful, Hossain, Rezaul Karim, Ruppa THulasiram, Neil D.B Bruce, Yang Wang (2018). "Hybrid Deep Learning Model

- for Stock Price Prediction". 2018 IEEE Symposium Series on Computational Intelligence (SSCI).
- [7] J.J. Wang, J. Z. Wang, Z. G. Zhang, and S. P Guo (2012). "Stock index forecasting based on a hybrid model". Omega, vol. 40, pp. 758-766.
- [8] C. Narendra Babu and B. Esvara Reddy (2014). "A moving-average filter-based hybrid ARIMA-ANN model for forecasting time series data". Applied Soft Computing, vol. 23, pp. 27-38.
- [9] Graves (2012). "Supervised Sequence Labelling with Recurrent". Studies in Computational Intelligence, Springer
- [10] Warbos, Paul J (1990). "Backpropagation through time: what it does and how to do it". Proceedings of the IEEE 78.10, pp.1550-1560.
- [11] C. Narendra Babu and B. Esvara Reddy, "Prediction of selected Indian stock using a partitioning-interpolation based ARIMAGARCH model," Applied Computing and Informatics, vol.11, pp. 130-143, July 2015.
- [12] R. K. Nayak, D. Mishra, and A. K. Rath, "A Naive SVM-KNN based stock market trend reversal analysis for Indian benchmark indices," Applied Soft Computing, vol. 35, pp. 670-680, October 2015.
- [13] C. H. Su and C. H. Cheng, "A hybrid fuzzy time series model based on ANFIS and integrated nonlinear feature selection method for forecasting stock," Neurocomputing, vol. 205, pp. 264-273, September 2016.
- [14] C. M. Anish and B. Majhi, "Hybrid nonlinear adaptive scheme for stock market prediction using feedback FLANN and factor analysis," Journal of the Korean Statistical Society, vol. 45, pp.64-76, March 2016.
- [15] L. Y. Wei, "A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting," Applied Soft Computing, vol.42, pp. 368-376, May 2016.
- [16] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches". arXiv preprint arXiv: 1409.1259,2014.
- [17] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model". Neurocomputing, vol. 50, pp. 159-175,2003.
- [18] Mohammad Obaidur Rahman, Md. Sabir Hossain\*, Ta-Seen Junaid, Md. Shafiul Alam Forhad, Muhammad Kamal Hossen, "Predicting Prices of Stock Market using Gated Recurrent Units (GRUs) Neural Networks", IJCSNS International Journal of Computer Science and Network Security, VOL.19 No.1,2019.
- [19] Bengio, Yoshua, S. Patrice, F. Paolo, "Learning long-term dependences with gradient descent is difficult". Neural Networks, IEEE Transactions on 5.2, pp.157-166,1994.
- [20] Cheng, Li-Chen, Yu-Hsiang Huang, and Mu-En Wu. "Applied attention-based LSTM neural networks in stock prediction." 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018.
- [21] Khare, Kaustubh, et al. "Short term stock price prediction using deep learning." 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). IEEE, 2017.
- [22] L. Di Persio and O. Honchar, "Artificial neural networks architectures for stock price prediction: Comparisons and applications," International Journal of Circuits, Systems and Signal Processing, vol. 10, pp. 403– 413, 2016.
- [23] M. F. Dixon, D. Klabjan, and J. H. Bang, "Classification-Based Financial Markets Prediction Using Deep Neural Networks," arXiv preprint, arXiv:1603.08604v2, June 2017.
- [24] Wenping Zhang, Chunping Li, Yunming Ye, Wenjie Li and Eric W.T. Ngai, "Dynamic Business Network Analysis for Correlated Stock Price Movement Prediction", IEEE Intelligent Systems Volume: 30, Issue: 2, Mar.-Apr. 2015.
- [25] Ouahilal, M., El Mohajir, M., Chahhou, M., & El Mohajir, B. E. "Optimizing stock market price prediction using a hybrid approach based on HP filter and support vector regression". 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt).
- [26] M. Khashei and M. Bijari (2011). "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting,". Applied Soft Computing, vol.11, pp. 2664-2675.
- [27] Mohammad Obaidur Rahman, Md. Sabir Hossain\*, Ta-Seen Junaid, Md. Shafiul Alam Forhad, Muhammad Kamal Hossen (2019). "Predicting Prices of Stock Market using Gated Recurrent Units (GRUs) Neural Networks". IJCSNS International Journal of Computer Science and Network Security, VOL.19 No.1.