

Application of Enhanced Hidden Markov Model in Stock Price Prediction

Donata D. Acula^{1,2}, Teofilo De Guzman¹



¹ Graduate School, Centro Escolar University, Manila

² Institute of Information and Computing Sciences, University of Santo Tomas, Manila

* Corresponding author email: ddacula@ust.edu.ph

Received: 25 May 2020 / Revised: 22 June 2020 / Accepted: 18 July 2020 / Published: 29 July 2020

ABSTRACT

The main focus of this research is the enhancement of the Hidden Markov Model by using some features of Neural Networks and the forecasted values of predictors by Seasonal Autoregressive Integrated Moving Average. The enhanced method was used to predict the close price of stocks whose predictors are open price, high price, low price, and volume of Apple and Nokia data. The performance of the method was measured using the Mean Absolute Percentage Error of the predicted price. The result was compared against the actual close price by using the paired T-test. The testing of the hypothesis showed that the Enhanced Hidden Markov Model obtained more than 94% accuracy rate. It also shows that in Apple data, the predicted close price of the Enhanced Hidden Markov Model is significantly better than the predicted close price of Neural Networks. Using Nokia data, the test claims that there is no difference between the performance of Enhanced Hidden Markov Model and Neural Network in prediction.

Keywords: Stock Price Prediction; Enhanced Hidden Markov Model; Neural Networks.

1 Introduction

Forecasting and prediction is now a current trend in the field of research. This is to optimize the resources or data available due to the advancement of technology. Various fields such as Computer Sciences, Medicine, Marketing, Economics and Business employed this type of researchers since most companies aimed to create an insight into their available data. In order to properly treat the data on hand, these companies spend a lot of resources. This is because they are fully aware that the extracted knowledge from the available data can be turned into the best decisions and solutions.

In the field of business, especially the stock market, hiring a good adviser or someone who will check the stock prices and trends is just an ordinary scenario. This is to study the chaotic system containing a large number of data and various factors that contribute to the price of the stock. A lot of researchers presented their study regarding the stock market prediction which is limited only in the index status. However, this paper created a forecast and prediction of prices specifically the close price which is more complicated because the predicted results are dependent on different variables or predictors.

Seasonal Autoregressive Integrated Moving Average (SARIMA) is one of the known techniques in forecasting in time series data with seasonality. SARIMA is a good model for prediction since its overall accuracy rate is more than 80% [1], [6], [11], [12], [15]. To improve the forecast and prediction, this paper used only the SARIMA in forecasting the predictors such as open price, low price, high price and volume. After employing the SARIMA in the forecasting of predictors, the Neural Networks were used to determine the predicted stock price known as Close price. Most of the papers mentioned above use the SARIMA

for forecasting the future values of the stocks without incorporating into another model especially the prediction models used in this study.

Neural Networks is one of the famous methods in predicting patterns when the relationship between inputs and outputs were complex. It has a robust ability to discover the relationship in the input data set without a priori assumption of the knowledge of the relation between the input and the output data and it can be used to build a model that identifies unknown hidden patterns in data which can be further used for prediction purposes [2], [10], [16]. The proposed method will be the combination of Seasonal Autoregressive Integrated Moving Average, Neural Networks and Hidden Markov Model. The developed model was presented in Section 5 of this paper and the model was called Enhanced Hidden Markov Model.

2 Seasonal Autoregressive Integrated Moving Average

Predicting a time series data is a challenging task [5]. In this paper, the training data underwent the forecasting using Seasonal Autoregressive Integrated Moving Average. The parameters $(p, d, q) = (2, 1, 2)$ and $(P, D, Q) = (0, 1, 1)$ for Apple data and $(p, d, q) = (0, 2, 1)$ and $(P, D, Q) = (0, 1, 1)$ for Nokia data where p is the number of autoregressive terms, d is the number of nonseasonal differences needed for stationarity, q is the number of lagged forecast errors in the prediction equation, P is the number of seasonal autoregressive terms, D is the number of seasonal differences and Q is the number of seasonal moving average [3],[7], [17].

After the simulation of the test, the summarized derived forecasting model for Apple and Nokia are the following:

$$\hat{Y}_t(\text{Apple}) = \mu + Y_t - 2\phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \theta_1 e_{t-1} - \Theta_1 e_t + \theta_1 \Theta_1 e_{t-1} \quad (1)$$

$$\hat{Y}_t(\text{Nokia}) = Y_t - 2\phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \theta_1 e_{t-1} - \Theta_1 e_t + \theta_1 \Theta_1 e_{t-1} \quad (2)$$

where

\hat{Y}_t = forecasted close price

μ = intercept

ϕ = coefficient of autoregressive

θ = coefficient of moving average

Θ = coefficient of seasonal moving average

Based on the equation of Apple as presented in (1), the derived equations for the predictors of Apple close price was shown in (3), (4), (5) and (6). The forecasting equation for Nokia Data was shown in (7), (8), (9) and (10) using (2).

$$\hat{Y}_t(\text{Open Price}) = 0.0000465 + Y_t - 2 + (0.89215)Y_{t-1} + (0.01475)Y_{t-2} - (1.55053)e_{t-1} - (0.99581)e_t + (-0.59327)(0.99581)e_{t-1} \quad (3)$$

$$\hat{Y}_t(\text{High Price}) = Y_t - 2(-0.39977)Y_{t-1} + (-0.55395)Y_{t-2} - (-0.50489)e_{t-1} - (0.99644)e_t + (-0.55573)(0.99644)e_{t-1} \quad (4)$$

$$\hat{Y}_t(\text{Low Price}) = Y_t - 2(0.00619)Y_{t-1} + (-0.34648)Y_{t-2} - (-0.06651)e_{t-1} - (0.99808)e_t + (-0.28512)(0.99808)e_{t-1} \quad (5)$$

$$\hat{Y}_t(\text{Volume}) = Y_t - 2(1.27347)Y_{t-1} + (-0.32416)Y_{t-2} - (1.78506)e_{t-1} - (0.99004)e_t + (-0.78722)(0.99004)e_{t-1} \quad (6)$$

$$\hat{Y}_t(\text{Open Price}) = -0.00076 + (-0.10333)e_{t-1} - (0.99282)e_t + (-0.0009067)(0.99282)e_{t-1} \quad (7)$$

$$\hat{Y}_t \text{ (High Price)} = -(-0.00152)e_{t-1} - (0.99528)e_t + (0.00836)(0.99528)e_{t-1} \quad (8)$$

$$\hat{Y}_t \text{ (Low Price)} = -(-0.001694)e_{t-1} - (0.99152)e_t + (0.07875)(0.99152)e_{t-1} \quad (9)$$

$$\hat{Y}_t \text{ (Volume)} = -(0.6457)e_{t-1} - (0.99895)e_t \quad (10)$$

The forecasted values of predictors for two weeks or ten days was shown in **Table 1** and **Table 2**.

The accuracy and error of the forecasting of predictors for both Apple and Nokia were presented in **Table 3** which was computed using Mean Absolute Percentage Error (MAPE) as shown in (11).

$$MAPE = 100 * \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right| \quad (11)$$

where:

A_i = the actual stock price

P_i = forecasted/predicted stock price

n = total number of test cases

Table 1: Forecasted Values of Predictors Using Apple Data

DATE	OPEN PRICE	HIGH PRICE	LOW PRICE	VOLUME
2018-12-31	153.380	158.553	154.343	43774812
2019-01-01	153.912	158.576	154.190	45752396
2019-01-02	153.898	158.458	154.203	47458648
2019-01-03	153.672	158.478	154.356	45231132
2019-01-04	153.387	158.462	154.325	48105639
2019-01-07	152.448	157.989	153.846	45183497
2019-01-08	153.040	157.920	154.058	45681188
2019-01-09	153.128	158.117	154.168	46930942
2019-01-10	152.994	158.063	154.194	44601811
2019-01-11	152.793	157.901	154.128	47494892

Table 2: Forecasted Values of Predictors Using Nokia Data

DATE	OPEN PRICE	HIGH PRICE	LOW PRICE	VOLUME
2018-12-31	5.6622	5.8203	5.6717	27380691
2019-01-01	5.6573	5.82	5.6538	29256318
2019-01-02	5.6618	5.8296	5.6495	28905019
2019-01-03	5.6615	5.8244	5.6378	31117286
2019-01-04	5.6608	5.8205	5.6426	28135446
2019-01-07	5.6506	5.8112	5.6463	27327953
2019-01-08	5.6455	5.8126	5.6431	29203580
2019-01-09	5.65	5.8222	5.6388	28852281
2019-01-10	5.6497	5.817	5.6271	31064548
2019-01-11	5.6489	5.8131	5.6319	28082708

Table 3: Error and Accuracy of the Forecasted Predictors using Apple and Nokia Data

Predictors	Error (%)		Accuracy (%)	
	Apple	Nokia	Apple	Nokia
Open	7.13487	2.07442	92.86513	92.86513
High	1.51459	1.86807	98.48541	98.48541
Low	1.66020	1.99107	98.3398	98.3398

3 Neural Networks

Neural Network is one of the best machine learning models for prediction and classification. It has a strong capability to determine the relationship in the input data set without a priori assumption of the knowledge of the relation between the input and the output data [10], [19]. It can be an aid to build a new model that identifies unknown hidden patterns in data sets [3],[8] .

The role of Neural Networks in this study is to create a prediction of the stock price given that the open, low, high price and volume was treated as predictors.

The weights produced by the networks were implemented in the Enhanced Hidden Markov model. The model chart of the Neural Networks for both Apple and Nokia data was shown in Figure 1 and Figure 2. These results were implemented using XLStat.

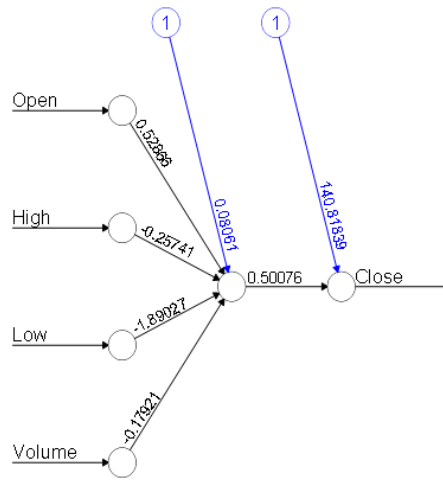


Figure 1: Model Chart of Neural Networks Using Apple Data

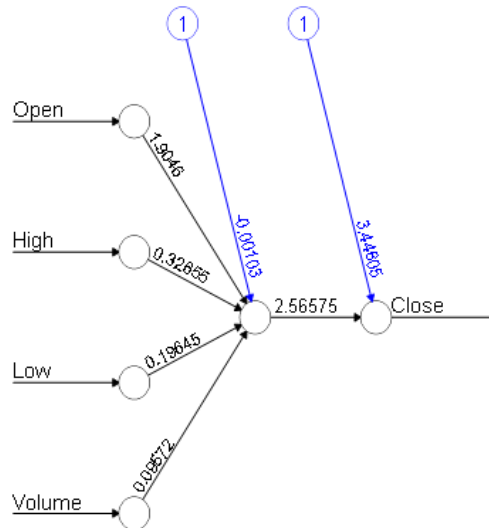


Figure 2: Model Chart of Neural Networks Using Nokia Data

4 Hidden Markov Model

A Hidden Markov Model consists of two stochastic processes. The first stochastic process is the Markov chain which is characterized by states and transition probabilities wherein the states were hidden. On the other hand, the second stochastic process produces emissions observable at each moment, depending on a state-dependent probability distribution [4], [13], [18], [20].

The development of Hidden Markov Models is widely used in various studies all over the world. Just like the Markov Models and Markov Chain Models, Hidden Markov Models are widely used also in prediction and analysis. This method applied in both one (1) dimensional such as voice recognition and enhancement and two (2) dimensional images.

In this study, the Hidden Markov Model was explored and enhanced by modifying the prediction stage of the model where the result of Seasonal Autoregressive Integrated Moving Average and Neural Networks were used. The enhanced model was implemented to predict the future close price of the stock of Apple and Nokia.

5 Prediction using Enhanced Hidden Markov Model

Figure 3 shows the complete process flow of the development of Enhance Hidden Markov Model. The gathered data from www.finance.yahoo.com of Apple and Nokia from 12-29-2014 to 01-11-2019 was converted into Time Series Data. The missing data were treated by getting the average price or volume per week. After the cleaning of data, it was divided into two sets, the training and testing set.

The training set underwent the forecasting of predictors using Seasonal Autoregressive Moving Average (SARIMA) which was discussed in **Section 2**. The same set of data was used to develop a Hidden Markov Model by calculating the Indicators, Average True Time Range and Logarithmic Returns respectively. The data frame for Hidden Markov Model was performed using the generated Average True Range and Logarithmic Returns. The generation of Hidden Markov Model took place using the Gaussian response distribution and the data frame and the model was fitted into the data sets. The Viterbi Path was obtained using the model and utilized the Baum Welch Algorithm to obtain the final matrix.

The y-intercept of the model was obtained by getting the multiplicative inverse of the trace of the matrix. The intercept and the weights obtained from the neural networks as discussed in Section III was used to create the enhanced model as shown in (12). The specific model for both Apple and Nokia Data was presented in (13) and (14).

$$\hat{y} = \frac{1}{2} \left(\left(\frac{1}{2} | a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 \log \log (x_4) | + b + c \right) + y_{NN} \right) \quad (12)$$

where

\hat{y} = predicted close price

a = weights of the predictors in Neural Networks

x = forecasted values of predictors from ARIMA

b = y-intercept of the model as computed using the Baum-Welch Algorithm

c = intercept to Neural Networks model

y_{NN} = close price intercept in Neural Networks model

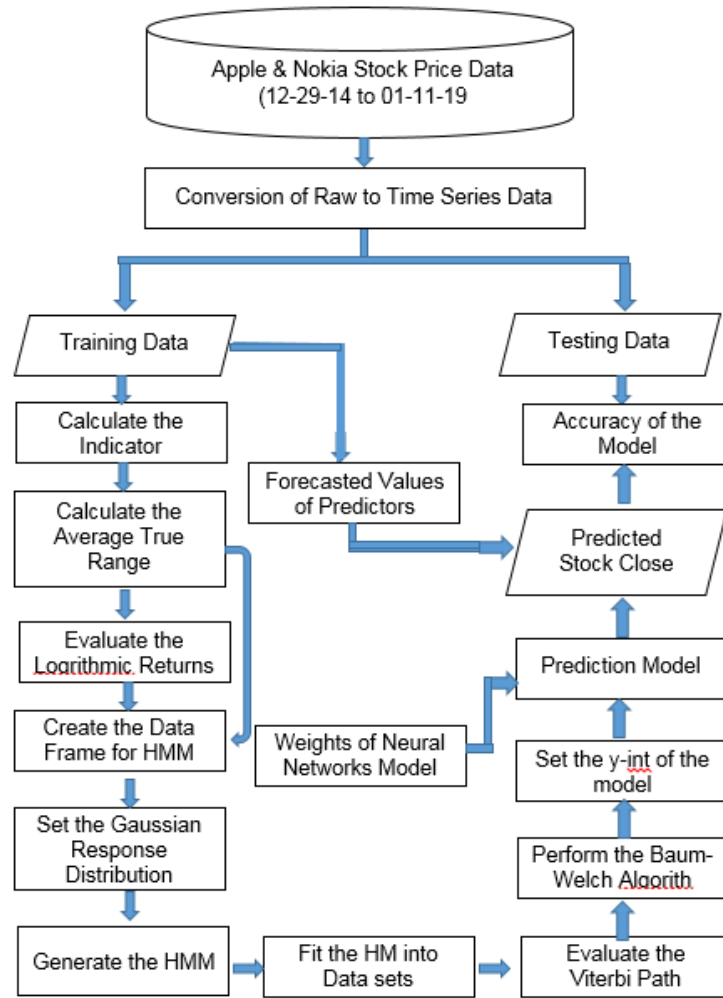


Figure 3: Process Flow of Prediction Using Enhanced Hidden Markov Model

$$\hat{y}_{APPLE} = \frac{1}{2} \left(\frac{1}{2} |0.52866x_1 - 0.52866x_2 - 1.89025x_3 - 0.17921 \log(x_4)| + 2.947413 + 0.08061 \right) + 140.81839 \quad (13)$$

$$\hat{y}_{NOKIA} = \frac{1}{2} \left(\frac{1}{2} |1.9046x_1 + 0.32655x_2 + 0.19645x_3 + 0.009572 \log(x_4)| + 2.947413 - 0.00103 \right) + 3.44605 \quad (14)$$

Using the model in (13) and (14), the results of the prediction of close price for two weeks for both Apple and Nokia data was shown in Table 4.

Table 4: The Predicted Close Price of Apple and Nokia Using Enhanced Hidden Markov Model

Date	Apple Close Price	Nokia Close Price
2018-12-31	144.061	5.95149
2019-01-01	143.9235	5.948356
2019-01-02	143.9175	5.951159
2019-01-03	144.0198	5.949902
2019-01-04	144.0434	5.949633
2019-01-07	143.876	5.944049
2019-01-08	143.8893	5.941604
2019-01-09	143.9569	5.944384
2019-01-10	143.9775	5.94312
2019-01-11	143.9547	5.942807

6 Model Evaluation

To evaluate the performance of the Enhanced Hidden Markov Model the error and accuracy of the model were computed using MAPE whose formula was shown in (11). To verify if there is no significant difference between the actual and predicted close price, the paired t-test was employed since both sets of data were normally distributed. The normality was performed using the Shapiro-Wilk Test which obtained a p-value of 0.7639 and 0.6636 for actual and predicted closed price of Apple. The Nokia data obtained a p-value of 0.587 and 0.1258 for actual and predicted close price respectively.

6.1 Error and Accuracy

As presented in **Section 5**, the predicted close price of Apple and Nokia was shown in **Table 4**, while **Table 5** shows the actual close price for the two data sets.

Table 5: The Actual Close Price of Apple and Nokia Using Enhanced Hidden Markov Model

Date	Apple Close Price	Nokia Close Price
2018-12-31	157.74	5.82
2019-01-01	149.4567	5.765
2019-01-02	157.92	5.74
2019-01-03	142.19	5.57
2019-01-04	148.26	5.93
2019-01-07	147.93	6.02
2019-01-08	150.75	6.15
2019-01-09	153.31	6.21
2019-01-10	153.8	6.14
2019-01-11	152.29	6.08

Table 6 shows the error and accuracy of predictions of Enhanced Hidden Markov Model for Apple and Nokia data.

Table 6. Error and Accuracy of the Predicted Close Price Using Apple Data

DATA SET	ERROR (%)	ACCURACY (%)
Apple	5.0624	94.9376
Nokia	3.0660	96.9340

6.2 Testing for Significant Difference

At 5% level of significance, it was hypothesized that there is no significant difference between the predicted and actual price. Using the paired t-test whose formula shown in (15). The result of testing was presented in **Table 7**.

$$t = \frac{\bar{d} - d_0}{\frac{s_d}{\sqrt{n}}} \quad (15)$$

where

t = the computed value of t;

\bar{d} = average of the deviation of the results between predicted and actual stock close price;

d_0 = assumed to be 0 since its H_0 claims that there is no significant difference between the predicted and actual price;

n = total number of test cases;

s_d = the standard deviation of the difference between the predicted and actual close price.

Table 7. The Result of Testing of Hypothesis for Apple Data

Data Set	p-value	Conclusion
Apple	0.0008263	There is a significant difference between the predicted and actual close price.
Nokia	0.9527	There is NO significant difference between the predicted and actual close price.

7 Conclusions

This paper used Seasonal Autoregressive Integrated Moving Average to forecast the predictors of close price. The forecasting models developed in this paper produced good metrics for accuracy which only implies that SARIMA is useful in dealing with the stocks data. The results of SARIMA were utilized in the Enhanced Hidden Markov Model to predict the close price of stock. Mean Absolute Percentage Error was used to measure the performance of the developed model and the outcome obtained a promising result for prediction with at least 95% accuracy. In order to verify if there is no difference between the actual and predicted close price, a paired t-test was implemented. Nokia Data concluded that at 5% level of significance, there is no significant difference between the predicted and actual close price. However, since there is a significant difference in the actual and predicted close price in the Apple Data, it is recommended to further explore the inputs of the Enhanced Hidden Markov Model. Future researchers are encouraged to use the developed model in other fields of applications.

8 Declarations

8.1 Competing Interests

Authors declare that no potential conflict of interest exists related to this article.

8.2 Acknowledgement

The researcher would like to thank the following who have guided and supported in the development of this Dissertation paper. To the panel of evaluators, Dr. Avelina R. Raqueño, Dr. Rene R. Belecina, Dr. Emma Lina F. Lopez, Dr. Evangeline F. Golla and Dr. Erna F. Yabut who gave their insights and recommendation to improve the manuscript. To the researchers' family, friends and colleagues, for the support and encouragement through the development of the research paper. And to the Lord, our God, for the knowledge, wisdom, determination and guidance to finish this study.

How to Cite this Article:

D. D. Acula and T. De Guzman, "Application of Enhanced Hidden Markov Model in Stock Price Prediction", *J. Mod. Sim. Mater.*, vol. 3, no. 1, pp. 70-78, Jul. 2020. <https://doi.org/10.21467/jmsm.3.1.70-78>

References

- [1] Mondal, P., Shit, L., & Goswami, S. (2014). Study of effectiveness of time series modeling (Arima) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, 4 (2), 13-29. DOI: <http://dx.doi.org/10.5121/ijcsea.2014.4202>
- [2] Bashambu, S., Sikka, A., & Negi, P. (2018). Stock Price Prediction Using Neural Networks. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4, 603-606. *Semantic Scholar*
- [3] Valipour, M. (2015). Long-term runoff study using SARIMA and ARIMA models in the United States. *Met. Apps*, 22: 592-598. doi:10.1002/met.1491
- [4] G.L.Kouemou "History and Theoretical Basics of Hidden Markov Models, Hidden Markov Models", Theory and Applications, Dr. Przemyslaw Dymarski (Ed.), ISBN: 978-953-307-208-1, InTech, doi: 10.5772/15205
- [5] Chindanur, narendra babu & B, Eswara. (2015). Performance comparison of four new ARIMA-ANN prediction models on internet traffic data. *Journal of Telecommunications and Information Technology*. 2015. 67-75.
- [6] Adhikari, Ratnadip & Agrawal, R.. (2013). An Introductory Study on Time series Modeling and Forecasting. 10.13140/2.1.2771.8084.
- [7] A. Adebisi, C. Ayo "Stock Price Prediction Using the ARIMA Model" Proceeding in 16th International Conference on Computer Modeling and Simulation, pp. 105-111, 2014.
- [8] Yawen Li, Weifeng Jiang, Liu Yang, Tian Wu, On neural networks and learning systems for business computing, *Neurocomputing*, Volume 275, 2018, Pages 1150-1159, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2017.09.054>.

- [9] J. Jagwani, M. Gupta, H. Sachdeva and A. Singhal, "Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend," *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2018, pp. 462-467, doi: [10.1109/ICCONS.2018.8663035](https://doi.org/10.1109/ICCONS.2018.8663035).
- [10] Nayak, S.C., Misra, B.B. & Behera, H.S. ACFLN: artificial chemical functional link network for prediction of stock market index. *Evolving Systems* 10, 567–592 (2019). <https://doi.org/10.1007/s12530-018-9221-4>
- [11] Yao Dong, Jianzhou Wang, He Jiang, Jie Wu, Short-term electricity price forecast based on the improved hybrid model, *Energy Conversion and Management*, Volume 52, Issues 8–9, 2011, Pages 2987-2995, ISSN 0196-8904, <https://doi.org/10.1016/j.enconman.2011.04.020>.
- [12] Yao Dong, Jianzhou Wang, He Jiang, Jie Wu, Short-term electricity price forecast based on the improved hybrid model, *Energy Conversion and Management*, Volume 52, Issues 8–9, 2011, Pages 2987-2995, ISSN 0196-8904, <https://doi.org/10.1016/j.enconman.2011.04.020>.
- [13] Pawar, R.V., Jalnekar, R.M. & Chitode, J.S. Review of various stages in speaker recognition system, performance measures and recognition toolkits. *Analog Integr Circ Sig Process* 94, 247–257 (2018). <https://doi.org/10.1007/s10470-017-1069-1>
- [14] Pedraza, L.F.; Hernandez, C.A.; Paez, I.P.; Ortiz, J.E.; Rodriguez-Colina, E. Linear Algorithms for Radioelectric Spectrum Forecast. *Algorithms* **2016**, *9*, 82. <https://doi.org/10.3390/a9040082>
- [15] W. Wang and Y. Guo, "Air Pollution PM2.5 Data Analysis in Los Angeles Long Beach with Seasonal ARIMA Model," 2009 International Conference on Energy and Environment Technology, Guilin, Guangxi, 2009, pp. 7-10, doi: [10.1109/ICEET.2009.468](https://doi.org/10.1109/ICEET.2009.468).
- [16] Dua, S., Sahni, S., Goyal, D. P.: Information Intelligence, Systems, Technology and Management. Proceedings of the 5th International Conference, ICISTM 2011, Gurgaon, India, March 10-12, 2011. ISBN 978-3-642-19423-8
- [17] Gelažanskas, Linas and Gamage, Kelum A. A., Forecasting Hot Water Consumption in Residential Houses, *Energies*, Vol. 8, 2015, No. 11, pp. 12702--12717, <https://www.mdpi.com/1996-1073/8/11/12336>
- [18] Luo, A.; Chen, S.; Xv, B. Enhanced Map-Matching Algorithm with a Hidden Markov Model for Mobile Phone Positioning. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 327. <https://doi.org/10.3390/ijgi6110327>
- [19] A. Victor Devadoss, T. Antony Alphonse Ligorì (2013). Forecasting of Stock Prices Using Multi Layer Perceptron. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.679.2370&rep=rep1&type=pdf>
- [20] M.sharmila mari, M.Ponnrajakumari (2014). Detection of Insider and Outsider Attack using Holistic Protocol in Vehicular Ad Hoc Networks. Retrieved from <http://ijaracet.org/wp-content/uploads/IJAR CET-VOL-3-ISSUE-3-972-975.pdf>

Publish your research article in AIJR journals-

- ❖ Online Submission and Tracking
- ❖ Peer-Reviewed
- ❖ Rapid decision
- ❖ Immediate Publication after acceptance
- ❖ Articles freely available online
- ❖ Retain full copyright of your article.

Submit your article at journals.ajjr.in

Publish your books with AIJR publisher-

- ❖ Publish with ISBN and DOI.
- ❖ Publish Thesis/Dissertation as a Book.
- ❖ Publish Monograph.
- ❖ Publish Edited Volume/ Book.
- ❖ Publish Conference Proceedings
- ❖ Retain full copyright of your books.

Submit your manuscript at books.ajjr.org