Mathematical model for implementing Non Linear Time Series Forecasting using Deep Learning

1Imran Qureshi, 2Burhanuddin Mohammad, 3Mohammed Abdul Habeeb, 4Mohammed Ali Shaik
1, 2, 3 Lecturer, IT Department, University of Technology and Applied Science-AlMusanna, Oman
4Assistant Professor, Department of Computer Science & Engineering, SR University, Warangal, Telangana, India

Email: imran@act.edu.om

Abstract. In this paper we propose a novel algorithm for implementing the mathematical model using the Non Linear Time Series Forecasting using the methods of deep learning which has to avoid noise and tends to transforms the space to next level in a iterative manner by swapping the down-trend to up-trend or up-trend to down-trend based on the iterative set values by computing the transformations. The proposed model must handle huge dimensional data which is in a dynamic nature of proposed model and then we need to compare the results attained from the proposed model with the complex ARIMA model and identified various trial and error methods for predicting the future price without guidance. For obtaining the results we used the petroleum dataset of India and attained the results which prove that our proposed mathematical model is far better than the famous ARIMA model while handling non linear time series applications.

Keywords: mathematical model, machine learning, forecasting, statistics, time-series

1. Introduction

In the present computer era that comprises of huge historic data Time Series (TS) [1] occurs anywhere the data is being acquired completely and clearly indexed based on the time aspect where as in a automated classification immense importance is attained in the field of petroleum data to assess the production and supply of petroleum products based on the market analysis the main focus is on comparing with the ARIMA model and perform classification to attain the decision based on the observations performed over partial samples. Classification of time series is implemented in the areas of health diagnoses, classification monitoring systems, predictive maintenance.

The mathematical model for implementing the TS classification represents the methodology of handing over a label or a specific class to the data set as there exist many implementations to deal with the TS classification issues as the time series classification utilizes only the available observations.

The basic Artificial Neural Network (ANN) [5] comprises of many samples as they comprises of connected processors called as the neurons where every neuron leads to production of series of real world scenarios that activate input neurons by which the environment is activated by which the hidden and output neurons will get activated by default using the weighted connections for attaining the output neurons by tuning the weights. This process is called as the learning model where each layer is
transformed into nonlinear model for activating and assigning the weights for the previous layer in an iterative manner based on the context into non-linear transformations [4].

2. Related work

In reference [2] authors have proposed a time series observation as xt where x denotes observation and t denotes time acquired in continuous and discrete time series implementation and gets normalized by using the range from 0 to 1 for a time interval T0 for attaining the time series observation based on the sampling rates acquired.

In reference [3] authors have proposed the performance aspect of Nearest Neighbor using the binary classification based on the labeled available data as the approach comprises of 1-NearestNeighbor classification over the mathematical unlabeled data in a iterative manner for accumulating the instances based on the classification over huge data set till a halting criteria is met.

In reference [4] authors have proposed the classification of 1-Nearest Neighbor with a univariate denoted as multivariate time series based on the distance measure acquired in 1-Nearest Neighbor algorithm which intern implements the Euclidean distance measure [5] and Frobenius distance [6] for attaining the multivariate time series by using the distinct warping windows using the optimized k-Nearest Neighbors search [7] over multivariate time series with distinct extensions of data till we attain a single value by constructing a covariance matrices for obtaining the equal dimensions in a multivariate time series matrices within the time series [8].

In reference [9] authors have proposed a Deep Neural Network [12] for handling the various implications of the computer vision that internally utilizes the tasks of the convolutional neural networks (CNN) [10] which has the properties such as sparse connectivity in which each and every layer is assigned with only one region in a input image considered as receptive fields, and shared weights as all layers have same set of weights or equal weights in a unsupervised context similar to Restricted Boltzmann Machines (RBM) [11] and Deep Belief Networks has multiple learning layers that are uniquely trained for every layer and Autoencoders are used to feed forward network design for replicating the mathematical models input [12].

3. ARIMA Model:

The ARIMA model comprises of tools and techniques for imparting the forecast or predict the implications of distinct TS data by identifying and sorting various conversions that are needed to stationarize the time series data as the ARIMA model that is implemented over the non regular data which can be further distinguished as an ARIMA(with parameters p,q,r) model in which the attribute p denotes total quantity attained with automatic regression elements then d represents total quantity of non recurring differentiation and q denotes total quantity attained for the predicted outliers generated in the below equation(1) which is required for prediction and the present behavior of any of the data values is depicted in provisions of linear associations with their existing possible data values.

The ARIMA model comprises of the Integrated Differencing Component (IDC) [3] where the implementation process comprises of various TS attained through inactive discrimination as the quantity of variability which is further implemented by indicating using d in the model which sequentially comprises of components: Auto Regressive (AR) [31] and Moving Average (MA) [32] as illustrated in below equation.

\[ Y_t \leftarrow (a_0) + (a_1 \cdot Y_{t-1}) + (a_2 \cdot Y_{t-2}) + \ldots + (-b_1 \cdot u_{t-1}) - (-b_2 \cdot u_{t-2}) - \ldots - (b_q \cdot u_{t-q}) + (u_t) \]  

(1)

The correlation attained relating to present value attained through time series based on a number of historical values is denoted by AR where as AR(1) denotes the association concerning the present value of t with its instantaneous precedent value attained during the time span (t-1) where the MA denotes the total time acquired to influence the random shock and MA(1) provides us with the correlation attained among the time interval t along with the shock at interval t-1 [5].
The major benefit of ARIMA models is optimality and comprehensiveness when we thoroughly analyze the family of models though it has its own limitations that are likely difficulty to identify the exact preferred model as it needs additional amount of time interval to do this and though some of the models can’t be interpreted structurally due to existence of outliers whose impact is clearly visible on the nonlinear time series data that poorly twists the verification and estimation process.

There existed many of the forecasting results models based on many of the researchers and most of the work is related to the model that implements implications on the past data because of the attributes of ARIMA model can’t be further tuned easily due to which a novel data has to be recreated completely from the scratch. And the periodic verification process has to be extensively applied periodically to validate the capability of ARIMA model that is applied or implemented for calculating and estimating ARIMA model with another model that is our proposed model though it is very typical to use as it requires good exposure and expertise in the area [6].

3.1 Proposed Mathematical Model

In this paper we proposed a model that is purely based on the implementation of nonlinear time series data with the stationarity transformations to perform forecasting as the variable behavior of time series data which is typical to be used directly into any of the application rather the time series data has to be transformed the data into stationarity information as the processes comprises of identification of Standard Deviation (SD) with a proportion of TS rate at a time period t and with the attained mean value is calculated using avg(t-1) where t-1 is the value at a given time period in a time series with latest maximum and minimum values of data given with the time span t-1 intend to compute the information. And the curve is plotted is yielded from the exponential value attained that corresponds to the equation 2 as the nonlinear time series data is further converted that by attaining the stationarity in terms of series is :

\[ P = Ae^{bt} \]  

(2)

In equation 2 values of A and b are unknown at the initial state as the equation is attained through exponential regression analysis and p value is attained by providing the SD at a given instance q and the equation used to attain the SD at a specific instance T+1 and s is required to resolve the anonymous data at T+1 value.

\[ \hat{x} = \frac{1}{n} \sum_{i=1}^{n} x_i^2 \]  

(3)

\[ SD = \left( \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{x})^2 \right)^{0.5} \]  

(4)

The SD equation will tend to diminish the quadratic equation which is in the form of \( ax^2 + bx + c = 0 \) with \( a = n(n + 1) \) and \( b = -2s(n + 1) \) and \( c = - (ns^2s) - (s^2s) + (m(n + 1)) * (nd * nd) * (n + 1)(n + 1) \), where the value of \( n+1 \) is denoted by the sum of all element values in time series based on the sum of the data and the value of m is attained by the aggregate of squares of data and d denotes the SD of \( n+1 \) attribute values where as the x is \( (n+1) \text{th} \) element that is available deliberately in time series since the outcome of x and the mean values when compared with the maximum and minimum values in a sequence values attained at a specific instance t-1 provides us with the probable attained value T+1.

Proposed Algorithm: the Predicting Mathematical model over Nonlinear TS data
Step1: Start
Step2: Initialize \( c_0 = \text{recent high value}, c_1 = \text{recent low value for every instance with the nonlinear time series data as it is significant to provide the negation of this is the downtrend or uptrend values considered for computing the values.} \)
Step3: identify averages of the ct which is time series data based on \( c_0 \) and \( c_1 \) values for each of the possible illustrations
Step4: identify and create proportion of TS data at specific to a given time span t with the possible mean value attained at a specific interval point t-1.
Step5: Generate the SD ratio based on the exponential decay curve form equation 2
Step 6: attain slope m
Step 7: attain the constant A by evaluating equation 2 using regression by substituting sum of y instead of sum of x
Step 8: calculate sum_of_products of (x, y)
Step 9: calculate sum_of_squares(x²)
Step 10: SD(x) = step 8 + step 9 at instance T+1
Step 11: implement equation 3, 4 over the unknown time series data at instance T+1
Step 12: calculate a = n*(n+1)
Step 13: calculate b = -(2*s)*(n+1)
Step 14: calculate c = -(n*s*n*s) - (s*s) + m * ((n+1) * (n+1)) – (nd * nd) * ((n+1)*(n+1))
Step 15: calculate s = sum_of_timeseries(a, b, c)
Step 16: calculate m = sum_of_squares(s)
Step 17: calculate SD = get_element(m, n+1)
Step 18: calculate x = average(high, low) at time t-1
Step 19: Stop.

By implementing the above algorithm future events are predicted using nonlinear time series data as the major difference between predicted value and the actual value is used to denote the feedback by which we can further reduce the error rate as frequently implemented in most of the forecasting models.

Sample Forecast

Let us now consider the example dataset considered in this paper is the daily oil consumption for the month of March 2020 by considering the dataset attained from https://www.ppac.gov.in/content/3_1_Petroleum.aspx to explain the proposed model where in figures 1(a) values 1 through 31 denotes the allotted time span from 03/01/2020 to 03/31/2020 and Figure 1(b) provides the definite revolutions of the day by day petroleum challenge in gallons for which the span as the attained conversions are foresee and implemented by using the equation y = Ae^(mx) which is further denoted in Figure 1(c) which illustrates the actual transformation by visualizing the higher prediction accuracy in the denoted transformations.

Figure 1. a) Daily petroleum demand in India for the month of March 2020 in gallons, b) Definite distorted series day by day petroleum demand for the month of March 2020 in gallons, c) Superimposition of predicted and actual series converted to day by day petroleum demand for the
month of March 2020 in gallons, d) Partial auto correlation of auto transformed series e) Actual and forecasted day by day petroleum required for the month of march 2020 in gallons

The serial dependencies of data that is transformed in series and examined based on the partial autocorrelation function which provides the total magnitude of association that defines the correlations as the partial autocorrelation transformed in series and the illustrated definite and foresee day by day petroleum required is illustrated in above figure 1.

The proposed model presents solution for predicting non-linear time series data as one of the multifarious verdict building essential by end users to design the ARIMA (with parameters p,d,q) model for predicting the accuracy which clearly implies on the collection of data values related to parameters as the anticipated model will not rely over any of the parameters as the different parameter values leads to distinct solutions and the model is free from the dependencies of parameters as it deliberates unique solution as it does not require much of the time to model the identification process which is required to predict the model to dynamically design the model by accommodating the changes with new values to handle high dimensional data which internally updates the design as it is based on the non-linear data models.

4. Experimental Results

For performing petroleum forecasting and its demand which is a important job to predict the custom demand based on the future demand for controlling the demand and price in the goods or commodity markets and the tentative association of the anticipated model with ARIMA model for forecasting the day by day petroleum required for the month of march 2020 in India is being utilized for the trailing of various data sources from: https://www.ppac.gov.in/content/3_1_Petroleum.aspx) from this source we have taken only historical data for experimentation and the performance is assessed based on the pre assigned metrics such as the Mean Square Error (MSE) [22] Root Mean Square Error (RMSE) [23] Mean Absolute Percentage Error (MAPE) [24] Prediction Accuracy (PA) [25] and overall time required to perform the aspects of prediction [29].

In order to do comparison of performance using the proposed mathematical model with ARIMA model is illustrated in Table 1 which shows deep learning algorithm with filters: MSE, RMSE and MAPE, the percentage increase in the prediction accuracy by increasing the time taken towards the proposed mathematical model by comparing with the ARIMA model.

Table 1. Comparison of performance for the proposed mathematical model with ARIMA model

| Deep Learning Filter | ARIMA Model | Proposed Model |
|----------------------|-------------|----------------|
| MSE                  | 1.62        | 1.42           |
| RMSE                 | 1.33        | 1.23           |
| MAPE                 | 1.45        | 1.32           |
| Predicted Accuracy   | 6.80%       | 4.90%          |
| Speedup              | 2.7         | 12.6           |
5. Conclusion:
In this paper we deliberately anticipated the unique algorithm for implementing a mathematical model using the non linear time series forecasting using the methods of deep learning that avoids noise and transforms the space to next level in a iterative manner that swaps the down-trend to up-trend or up-trend to down-trend based on the iterative set values while computing the transformations as per the anticipated algorithm. As the anticipated model has the capability to handle effectively the complex aspects of component data which is in a vigorous nature of the anticipated model and then we compared the results attained as per the anticipated model with the complex ARIMA model and identified various trial and error methods and implemented over the anticipated algorithm is capable of
predicting future price without expert guidance. In experimental study we used the petroleum dataset of India and attained the results which prove that our anticipated mathematical model represents to be far better than the famous ARIMA model while handling non linear time series applications.

References
[1] Uma, S., A. Chitra, and J. Suganthi 2013 Design of a non-linear time series prediction model for daily petroleum demand forecasting International Journal of Business Innovation and Research, pp 1-5
[2] N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny.2010 An empirical comparison of machine learning models for time series forecasting Econometric Reviews vol 29(5-6) pp 94-621
[3] Z. Banko and J. Abonyi 2012 Correlation based dynamic time warping of multivariate time series Expert Systems with Applications vol 39(17) pp12814-12823
[4] B. Batres, Estrada 2015Deep learning for multivariate financial time series Expert Systems with Applications vol 59(16) pp12824-12833
[5] Y. Bengio 2009 Learning deep architectures for ai Foundations and trends in Machine Learning vol 2(1) pp 1-127 vol 43 pp122-133
[6] P. Brockwell and R. Davis 2013 Time Series: Theory and Methods Springer Series in Statistics. (Springer NewYork).
[7] E. Busseti, I. Osband, and S. Wong 2012 Deep learning for time series modeling Technical report, Technical report,(Stanford University)
[8] K. Chakraborty, K. Mehrotra, C. K. Mohan, and S. Ranka 1992 Forecasting the behavior of multivariate time series using neural networks Neural networks vol 5(6) pp 961-970
[9] Y. Chen, E. Keogh, B. Hu, N. Begum, A. Bagnall, A. Mueen, and G. Batista 2015 The ucr time series classification (archive), July
[10] M. Langkvist, L. Karlsson, and A. Lout 2014 A review of unsupervised feature learning and deep learning for time series modeling Pattern Recognition Letters vol 42 pp 11-24
[11] Mohammed Ali Shaik, P. Praveen 2020 A Survey On Various Clustering Algorithms Using Nature Inspired Algorithms Journal Of Mechanics Of Continua And Mathematical Sciences Vol 15(7) pp 282-293
[12] Zhang, G 2003 Time series forecasting using a hybrid ARIMA and neural network model Neurocomputing vol. 50 pp 159-175
[13] Ali Shaik, M 2020 Time series forecasting using vector quantization International Journal of Advanced Science and Technology vol 29(4) pp. 169-175
[14] Makridakis, S., Spiliotis, E., and Assimakopoulos, V 2018 Statistical and Machine Learning forecasting methods: Concerns and ways forward PloS one vol 13(3) e0194889.
[15] Shaik, M.A 2019 A survey on text classification methods through machine learning methods International Journal of Control and Automation vol 12(6) pp 390-396
[16] Box, G.E.P. and Jenkins, G 1990 Time Series Analysis, Forecasting and Control Holden-Day Incorporated
[17] Shaik, M.A., Sampath Kumar, T., Praveen, P., Vijayaprakash, R 2019 Research on multi-agent experiment in clustering International Journal of Recent Technology and Engineering vol 8(1 Special Issue 4) pp. 1126-1129
[18] Holt, C.E 2004 Forecasting seasonals and trends by exponentially weighted averages International Journal of Forecasting vol 20(1) pp 5-10
[19] Winters, P.R 1960 Forecasting sales by exponentially weighted moving averages Management Science vol 6 pp 324–342.
[20] hengpingChe, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu 2016 Recurrent neural networks for multivariate time series with missing values arXiv preprint arXiv:1606.01865
[21] Mohammed Ali Shaik 2019 A Survey of Multi-Agent Management Systems for Time Series
Data Prediction International Journal of Grid and Distributed Computing (IJGDC) Vol 12(3) pp 166-171

[22] P. Praveen, C. J. Babu and B. Rama 2016 Big data environment for geospatial data analysis International Conference on Communication and Electronics Systems (ICCES)(CESYS.2016.7898816) pp 1-6

[23] Wang, Z., W. Yan, and T. Oates 2017 Time series classification from scratch with deep neural networks: A strong baseline In Neural Networks (IJCNN) International Joint Conference USA IEEE pp 1578–85

[24] Weisstein, E. W 1999 Convolution. Accessed November 30, 2018. http://mathworld.wolfram.com/Convolution.html.

[25] Wiens, J., E. Horvitz, and J. V. Guttag 2012 Patient risk stratification for hospital-associated c. diff as a time-series classification task In Advances in Neural Information Processing Systems (Curran Associates Inc) pp 467–75

[26] Xing, Z., J. Pei, and E. Keogh 2010 A brief survey on sequence classification ACM Sigkdd Explorations Newsletter vol 12(1) pp 40–48

[27] Shaik, M.A 2018 Protecting Agents from Malicious Hosts using Trusted Platform Modules (TPM) Proceedings of the International Conference on Inventive Communication and Computational Technologies, (ICICCT 2018) pp. 559-564

[28] Zhao, B., L. Huazhang, S. Chen, J. Liu, and W. Dongya 2017 Convolutional neural networks for time series classification Journal of Systems Engineering and Electronics vol 28(1) pp 162–69

[29] Praveen P., Rama B 2018 A Novel Approach to Improve the Performance of Divisive Clustering-BST In: Satapathy S., Bhatija V., Raju K., Janakiramaiah B. (eds) Data Engineering and Intelligent Computing, Advances in Intelligent Systems and Computing (Springer, Singapore) vol 542

[30] Binkowski M, Marti G, Donnat P 2018 Autoregressive Convolutional Neural Networks for Asynchronous Time Series Proceedings of the International Conference on Machine Learning (ICML)

[31] Mohammed Ali Shaik, P.Praveen, Dr.R.Vijaya Prakash 2019 Novel Classification Scheme for Multi Agents Asian Journal of Computer Science and Technology Vol.8 (S3) pp. 54-58

[32] Moraffah R, Karami M, Guo R, Raglin A, Liu H 2020 Causal Interpretability for Machine Learning–Problems Methods and Evaluation. arXiv e-prints. arXiv:2003.03934.