Enhanced ANN training model to smooth and time series forecast

Mohammed Ali Shaik¹ and Dhanraj Verma²
¹Department of Computer Science & Engineering, SR University, Warangal, Telangana, India
²Professor, Department of Computer Science & Engineering, Dr. APJ Abdul Kalam University, Indore, India

Email: niharali@gmail.com

Abstract. In the present era most of the researchers has done extensive research in forecasting the time series data using distinct “linear, nonlinear and hybrid linear models” that combines models like “Auto Regressive Integrated Moving Average (ARIMA)” using “Artificial Neural Networks (ANN)” by combining the “Time Series Forecasting (TSF)”. In this paper we proposed a new algorithm which will smooth the data and evaluate the error rate to obtain time series predictions by considering Telangana state rainfall dataset. Our experimental illustrations propose that the results attained are statistically promising with the datasets that are used for generating the acceptable prediction result set.

Keywords: ARIMA, machine learning, statistics, time-series, ANN, TSF

1. Introduction
In the present era of internet forecasting model is a emerging tool to identify the relationship that exists in between the past data for attaining the future predicted values using either model driven or data driven approach. In model driven forecasting approach researchers possessing deep understanding of underlying data choose a model that is based on scientifically established relationships followed by forecasting [1]. The identification process of mathematical function or model is very typical once as the “Data Generating Processes (DGP)” generates the time series data based on the initially specified model for obtaining the appropriate methodology to capture the complex and dynamic behavior of time series data that is driven by the forecasting models for determining model parameters that predict the future values. The process of identification of underlying complex patterns of time series data requires a model that is built upon data to glorify chances to capture complex patterns of time series data by performing robust mining using the learning techniques to learn the underlying patterns economically.

The “Time series forecasting (TSF)” is the predominant process of forecasting the upcoming conclusions of a observable fact by analytically analyzing its earlier observations by computing the value of $Y_{n+h}$ and by analyzing the time series data $y = [y_1, y_2, ..., y_n]^T$ is known as TSF where each $y_i$ represents the forecasting horizon and is also called as step-ahead forecasting and if it leads to more than one hence it is called as a multi step-ahead forecasting method. In this paper we consider the one step-ahead forecasting method that achieves multi step-ahead forecasting by iterating the one step-ahead forecasting in a repeated manner.

Traditionally, statistical model such as “Auto Regressive Integrated Moving Average (ARIMA)” is used for performing the exponential smoothing which has been most often used in TSF due to its ease
in understanding by implementing these models as the statistical models assume linear correlation structure in time series fails to capture the nonlinear patterns efficiently [2]. Nonlinear statistical models like “Auto Regressive Conditional Heteroskedasticity (ARCH)” and “Generalized ARCH (GARCH)” have been used to model nonlinear time series as both the ARCH and GARCH family has thousands of models with each model that is capable of capturing a specific type of nonlinear pattern efficiently as it becomes complicated to select a suitable model without understanding the characteristics of time series.

The data driven “Machine Learning (ML) models” have the capability to extract data in an iterative manner for attaining the probable required patterns that exist in data by which comprises of robust data which initially specifies the model where the data has to identify the predominant model and its distinct attributes where as in ML this identification process is unknown as it improves the chances of capturing the underlying complex relationships that are available in the time series. And ANN model has shown promising results in performing forecasting competitions using M3 and NN3 that provides acceptable results using statistical models for TSF [3].

The “Artificial Neural Networks (ANN)” are the adaptive systems that tend to change its structure based on the information that is presented to it using nonlinear data that is driven by the mathematical models that do not require any a initial specific assumptions that are related to data by implementing complex relationships that exists between input(s) and output(s) that are based on the learning algorithms such as “supervised, unsupervised, semi-supervised or reinforcement”.

2. Literature review

2.1. ANN based forecasting methods:
The TSF is one of the key fields of research in statistic, management studies and now in computer science where efficient TSF have significant practical value and widespread usage in many fields and has performed predominantly using statistical based methods and due to unprecedented growth in computational power of machines and availability of variety of time series and ML techniques like ANN have become one of the forefront models in TSF [2]. When compared to traditional statistical techniques ML techniques like ANN possess several distinctive characteristics such as “nonparametric, data driven, nonlinear, flexible models and thus leads to greater capability to capture various underlying complex relationships existing in time series though TSF methods based on machine learning techniques that have been developed and applied in diversified application areas” [3]. Similarly ANNs are the one of the most popular classes that is based on the ML models which has predominantly gained huge attention in TSF which avoids distinct such models [4].

Over the last few decades in this research area various class of ANN models such as “Feed Forward Neural Networks” [5], “Recurrent Neural Networks” [6], “Single Layered Neural Networks” [7], “Multilayer Neural Networks” [8], “Real Valued Neural Networks” [9] and “Complex Valued Neural Networks” [10] comprising a variety of ANN models viz. “Multilayer Perceptron (MLP)” [11], “Deep Belief Network” [12], “Adaline” [13], “Stochastic Neural Network” [14], “Extreme Learning Machine” [15], “Radial Basis Function Networks (RBFN)” [16], “Beta Basis Function Neural Network (BBFN)” [17], “Generalized Regression Neural Network (GRNN)” [18], “Functional Link Artificial Neural Network” [19] etc. have been extensively used in TSF.

To improvise the accuracy of forecasting ANN based forecasting methods there exists several factors which have been taken into consideration in different studies where the reference [20] presents the issues affecting the implications of ANN based forecasting methods that includes factors such as: (a) “Activation or Transfer Function” [21], (b) “Design of ANN architecture” [22], (c) “Training algorithm” [23], (d) “Pre-processing techniques” [24].

Reference [21] claims the “process of activation function that plays a vital role in the convergence of training algorithms and have a significant effect on the performance of ANN based forecasting methods using binary sigmoid activation function at hidden layer and linear activation function at
output layer has been widely used in ANN based TSF methods for identifying the most appropriate activation function in ANN for financial TSF”.

Reference [22] claims the “performance of ANN based forecasting architecture that designs appropriate ANN architecture for determining the number of inputs using hidden layers called as neurons in each hidden layer and outputs using TSF that represents the past data points of the time series which is used to identify the unrevealed nonlinear relationship between data and using number of inputs in neural network modelling does not affect the bias of the model but affects the model variance significantly”.

Reference [23] claims the “training algorithms are used to determine the performance of features and based on which it selects the features that are selected based on the significant number of inputs using a filter approach based on the time resolution of data by developing another filter approach by using the ACF and PACF of the series to determine the number of inputs of ANN for univariate TSF. Once the number of inputs of ANN is selected by determining the number of hidden layers nodes plays a vital role in the performance of ANN based TSF methods as the neurons are present in hidden layers that allows neural network to non-linearly map the inputs with outputs by allowing the neural network to capture the underlying patterns”[48].

Reference [24] claims the “Pre-processing of time series for ANN based forecasting methods include treatment of time series components like trend and seasonality and normalization of time series and the previous studies determine the mixed results relating to ANN models alone are capable to handle trend component of time series or not by de-trending time series by taking the first difference is the best practical way for developing ANN based forecasting methods by modeling time series using ANN models are insensitive to de-trending by taking the first difference of the series and ANN models alone are quite capable to handle the trend components that are not sensitive to de-personalization by subtracting the seasonal average method by normalization technique that affects the performance of ANN based forecasting methods significantly”.

2.2. Hybrid ANN and statistical forecasting method

Reference [25] claims the “approach of hybridizing linear models with nonlinear models for TSF by assuming the time series data as a summation of a nonlinear series and a linear series as the ARIMA model is directly fitted to the given time series data to capture the linear patterns existing in the time series where the forecasts are subtracted from the original time series data to obtain the residual error series by considering the residual error series as nonlinear to obtain the forecasts as the final forecasts are computed by adding forecasts on linear series with forecasts on nonlinear component forecasts by proposing a hybrid ARIMA-ANN model trained with an optimized conjugated training algorithm and applied it to forecast the water quality time series data”.

Reference [26] claims the “ARIMA model is fitted to the given time series to capture the linear patterns though ARIMA model provide a better fit when a stationary time series is Gaussian in nature as the ARIMA model coefficients are computed using GMLE procedure which assumes the time series to be gaussian in nature as the application of ARIMA to a non-Gaussian time series results in poor fitting of ARIMA model and thus negatively affect the performance of a fore mentioned hybrid models as a whole”.

Reference [27] claims the “time series data into low-volatile and high-volatile series using a MA filter where the MA filter length is fine tuned in such a way that one series have a kurtosis value equal to the low volatility series and the other have a higher/lower kurtosis value than the constant value 3. Then the time series data is divided into low volatile and high volatile series as the resultant series is modelled by using ARIMA model and ANN model respectively after which the final forecasts are calculated by adding the ARIMA forecasts with ANN forecasts as the hybrid model considering three time series datasets using an exponential smoothing filter and further divided the high-volatile series into linear series and nonlinear series”[47].

Reference [28] claims the “superior performance of hybrid models than individual statistical and machine learning models are attained based on the systematic study that has been made to evaluate the
effectiveness of the hybrid models considering a variety of time series data using the ARIMA whose effectiveness of ETS in hybrid forecasting models have not yet been tested though both the models have their own strengths and weaknesses and can be used complementarily to cover up the weaknesses of each other”.

2.3. Fuzzy time series forecasting using ANN

By the development of forecasting methods based on fuzzy set theory many researchers have done research and found the following recommendations:

Reference [29] claims the “Identifying the universe of discourse (UOD) and partitioning the UOD into number of intervals that is dividing into several partitions with each representing a fuzzy set with UOD and the time series is calculated by taking a realistic decrement and increment on the minimum and maximum value of the series with the attributes such as high trend, the slope of the trend line and forecast horizon can be taken into consideration so that the future values will fall within the limits of UOD”.

Reference [30] claims the “Defining the fuzzy sets and fuzzifying the time series data by identifying the time intervals in terms of obtained time series observations that are fuzzified by associating a membership value of the interval having highest degree of membership for that observation”.

Reference [31] claims the “establishing the process of fuzzy logical relationships (FLR) where the order of the model is based on FLR value is defined on the order of the model using rule based methods to reduce the computational complexity and improve the forecasting accuracy using the multilayer feed forward neural network, multiplicative neuron model, fuzzy inference system, PSO, rough set theory have been used as an alternative to define FLR which has gained tremendous attention”[45].

Reference [32] claims the “forecasting and defuzzifying the forecasted values using the mid-point of intervals, interval average, centroid based methods in the defuzzification step, though the defuzzification by using interval average provides superior forecasting accuracy than other defuzzification alternatives at the cost of higher computational overhead”[46].

3. Proposed methodology

The performance of any ANN based solution clearly depends on the algorithm that is used for its training and iteratively update the weights of ANN for minimizing the error function which is defined as the desired and target output.

The algorithm is incorporated by the aspect of multiple trial vectors generated for each target vector with the best fitness to perform crossover operation to generate a trial vector and a target vector to produce the chromosomes for the next generation based on the obtained rate of change and the fitness of the best chromosome in the ith generation and later generates multiple vectors[49-51].

Algorithm: proposed Training Algorithm

Input: num_of_nurons, pattern_set, num_trans, validation_patterns
Output: predicted list
Step 1: Initialize all input parameters along with counter
Step 2: Randomly initialize population with genes representing weight of ANN
Step 3: evaluate fitness value of each data set attribute value
Step 4: generate vectors with vertical scaling factor
Step 5: generate crossover probability with horizontal scaling
Step 6: select trivial vector with best fit values
Step 7: update searching value based on the input criteria
Step 8: goto step 3 till each data value in data set is evaluated
Step 9: generate prediction result set with best fitness and optimal weights
Step 10: stop

The above algorithm can be represented graphically in the figure 1.
4. Experimental results
The time series data set of Telangana rainfall is considered to execute the above specified algorithm and the sample dataset snap shot is:

**Figure 1.** Flow diagram of proposed training algorithm

**Figure 2.** Dataset sample screen shot

The results attained after implementing the algorithm over district rainfall dataset:
5. Conclusion

In this paper we have done extensive research in the areas of forecasting Time Series data that utilizes distinct “linear or nonlinear and hybrid linear models” that combines linear models like Auto Regressive Integrated Moving Average (ARIMA) model using Artificial Neural Networks (ANN) by combining the Time Series Forecasting TSF. And we also proposed a new algorithm which will smooth the data and evaluate the error rate to obtain time series predictions by considering Telangana state rainfall dataset and we conclude that the attained experimental results denotes that the proposed model in this paper is statistically promising and generates acceptable prediction result set. In the future scope we would like to implement this algorithm using multi agent system.

6. References

[1] Babu, C.N., Reddy, B.E. 2014 A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data. Appl. Soft Comput vol 23 pp 27–38

[2] Behera, H.S., Dash, P.K., Biswal, B., 2010. Power quality time series data mining using S-
transform and fuzzy expert system. Appl. Soft Comput vol 10 (3) pp 945–955

[3] Box, G.E.P., Jenkins, G., 1990. Time Series Analysis, Forecasting and Control. Holden-Day Incorporated.

[4] Chandra, D.R., Kumari, M.S., Sydulu, M., 2013. A detailed literature review on wind forecasting International Conference on Power, Energy and Control, ICPEC pp. 630-634.

[5] Demsar, J., 2006. Statistical comparisons of classifiers over multiple datasets. J. Mach. Learn. Res vol 7 pp1–30

[6] de Oliveira, J.F., Ludermir, T.B., 2016. A hybrid evolutionary decomposition system for time series forecasting Neurocomputing vol 180 pp 27–34

[7] Donkor, E., Mazzuchi, T., Soyer, R., Alan Roberson, J., 2014 Urban water demand forecasting: review of methods and models. J. Water Resour. Plann. Manage. Vol 140 pp 146–159

[8] Enders, W., 1995 Applied Econometric Time Series, first ed. Wiley, New York.

[9] Hamilton, J.D., 1994. Time Series Analysis, first ed Princeton University Press, Princeton.

[10] Hollander, M., Wolfe, D.A., Chicken, E., 1999. Nonparametric Statistical Methods. John Wiley & Sons, Hoboken, NJ.

[11] Hyndman, R.J., TimeSeriesDataLibrary, 2010. URL http://data.is/TSDLdemo.

[12] Hyndman, R.J., Khandakar, Y., 2008 Automatic time series forecasting: the forecast package for R J. Statist. Softw vol 27(3) pp 1–22

[13] Hyndman, R.J., Koehler, A.B., 2006 Another look at measures of forecast accuracy Int. J. Forecast vol 22(4) pp 679–688

[14] 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

[15] Ali Shaik M.2020 Time series forecasting using vector quantization. Int J Adv Sci Technol vol 29(4) pp 169-175

[16] Mohammed Ali Shaik 2019 A Survey on Text Classification methods through Machine Learning Methods International Journal of Control and Automation (IJCA) vol 12(6) pp 390-396

[17] Box, G. Jenkins, G 1970 Time Series Analysis Forecasting and Control/”Holden Day, San Francisco, California John Wiley & Sons: Hoboken, NJ, USA

[18] Mohammed Ali Shaik, T. Sampath Kumar, P. Praveen, R. Vijayaparaksh 2019 Research on Multi-Agent Experiment in Clustering International Journal of Recent Technology and Engineering (IJRTE) vol 8(154) pp 1126-1129

[19] Chatfield, C 1978 The Holt–Winters forecasting procedure. J. R. Stat. Soc. Ser vol 27 pp 264–279

[20] De Livera, A.M. Hyndman, R.J. Snyder, R.D. 2011 Forecasting time series with complex seasonal patterns using exponential smoothing J. Am. Stat. Assoc vol 106 pp 1513–1527

[21] Mohammed Ali Shaik 2020 Time Series Forecasting using Vector quantization International Journal of Advanced Science and Technology (IJAST) vol 29(4) pp 169-175

[22] Salinas, D. Flunkert, V. Gasthaus, J. Januschowski, T 2019 DeepAR: Probabilistic forecasting with autoregressive recurrent networks Int. J. Forecast vol 20 pp 19-25

[23] 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

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

[25] Kumar RR, Reddy MB, Praveen P 2019Text classification performance analysis on machine learning Int. J Adv Sci Technol vol 28(20) pp 691-697

[26] Ravi Kumar R, Babu Reddy M, Praveen P 2019 An evaluation of feature selection algorithms in machine learning Int J Sci Technol Res vol 8(12) pp 2071-2074

[27] Praveen P, Jayanth Babu C 2019 Big Data Clustering: Applying Conventional Data Mining
Techniques in Big Data Environment *Lect Notes Networks Syst* vol 74 pp 509-516

[28] Mounika S, Sampath Kumar T 2019 Client –requirement fulfil QOS in multi server for max beneficial in cloud computing *Int J Adv Sci Technol* vol 28(7) pp 44-49

[29] Sampath Kumar T, Manjula B, Srinivas D 2017 A new technique to secure data over cloud *J Adv Res Dyn Control Syst* vol 2017(11) pp 391-396

[30] Ravi Chythanya K, Sunil G, Sudheer Kumar K, Korra SN, Harshavardhan A 2019 Security and safety in amazon EC2 service–A research on EC2 service AMIs *Int J Innov Technol Explor Eng* vol 8(4) pp 736-738

[31] Mohmmad S, Sheshikala M, Shabana 2018 Software defined security (SDSec): Reliable centralized security system to decentralized applications in SDN and their challenges *J Adv Res Dyn Control Syst* vol 10(10) pp 147-152

[32] Naresh Kumar S, Pramod Kumar P, Sandeep C, Thirupathi V, Shwetha S. 2019 A study on deep Q-learning and single stream Q-network architecture *Int J Adv Sci Technol* vol 28 pp 586-592

[33] Pramod Kumar P, Sandeep CH, Naresh Kumar S. 2018 An overview of the factors affecting handovers and effective highlights of handover techniques for next generation wireless networks *Indian J Public Health Res Dev* vol 9(11) pp 722-725

[34] Harshavardhan A, Babu S, Venugopal T. 2017 Brain tumor segmentation methods-a survey *J Adv Res Dyn Control Syst* vol 9(11) pp 240-245

[35] Kothandaraman D, Sheshikala M, Seena Naik K, Chanti Y, Vijaykumar B. 2019 Design of an optimized multicast routing algorithm for internet of things *Int J Recent Technol Eng* vol 8(2) pp 4048-4053

[36] Chanti Y, Raman K, Seenanaik K, Mahesh D, Bhaskar B. An enhanced on bidirectional li-fi attocell access point slicing and virtualization using das2 conspire. *Int J Recent Technol Eng* 20198(2 Special Issue 3):1109-1120.

[37] Chanti Y, Raman K, Seenanaik K, Mahesh D, Bhaskar B 2019 An enhanced on bidirectional li-fi attocell access point slicing and virtualization using das2 conspire. *Int J Recent Technol Eng* vol 8(Special Issue 3) pp 1109-1120

[38] Chythanya KR, Kumar KS, Rajesh M, Tharun Reddy S 2020 Sensor Cloud: A Breakdown information on the Utilization of Wireless Sensor Network by Means of Cloud Computing. *Test Eng Manage* vol 82 pp 13945-13954

[39] Shaik MA 2019 A survey on text classification methods through machine learning methods *Int J Control Autom* vol 12(6) pp 390-396

[40] Vijaya Prakash R, Taduri S 2020 Safe Navigation for Elderly and Visually Impaired People Using Adhesive Tactile Walking Surface Indicators in Home Environment. *Adv Intell Sys Comput* vol 933 pp 771-778

[41] Pasha SN, Harshavardhan A, Ramesh D, Md S, Shabana 2019 Variation analysis of artificial intelligence, machine learning and advantages of deep architectures. *Int J Adv Sci Technol* vol 28(17) pp 488-495

[42] Sandeep CH, Naresh Kumar S, Pramod Kumar P 2018 Security challenges and issues of the IoT system. *Indian J Public Health Res Dev* vol 9(11) pp 748-753

[43] Binkowski M, Marti G, Donnat P 2018 Autoregressive Convolutional Neural Networks for Asynchronous Time Series. In: *Proceedings of the International Conference on MachineLearning (ICML)* pp 2018-2025.

[44] Jamalpur, Bhavana 2020 Data Exploration As A Process Of Knowledge Finding And The Role Of Mining Data Towards Information Security *Journal Of Mechanics Of Continua And Mathematical Sciences* vol 15(6) 223-245

[45] Harshavardhan A, Suresh Babu Dr and Venugopal T Dr 2017 “Brain tumor segmentation methods – A Survey” *Jour of Adv Research in Dynamical & Control Systems* 11 240-245

[46] Harshavardhan A , Suresh Babu Dr and Venugopal T Dr 2016 “3D Surface Measurement through Easy-snap Phase Shift Fringe Projection”, *Springer conference International*
Conference on Advanced Computing and Intelligent Engineering Proceedings of ICACIE 1 179-186

[47] Mahender K, Ramesh KS and Kumar TA 2017 An efficient ofdm system with reduced papr for combating multipath fading Journal of Advanced Research in Dynamical and Control Systems 9(Special issue 14) 1939-1948

[48] Harshavardhan A, Suresh Babu and Dr, Venugopal T Dr 2017 “An Improved Brain Tumor Segmentation Method from MRI Brain Images” 2017 2nd International Conference On Emerging Computation and Information Technologies (ICECIT) IEEE 1–7. DOI.org (Crossref) doi:10.1109/ICECIT.2017.8453435.

[49] Rajasri I, Gupta AVSSKS and Rao YVD 2014 Effect of Symmetry on Structures of Epicyclic Gear Trains International Journal of Engineering and Technology Innovation 4(2) 77-85

[50] A.Harshavardhan Syed Nawaz Pasha Sallauddin Md D Ramesh 2019 “Techniques used for clustering data and integrating cluster analysis within mathematical programming” journal of mechanics of continua and mathematical sciences 14(6) 546-57, DOI.org (Crossref)https://doi.org/10.26782/jmcms.2019.12.00038

[51] Seena Naik K and Sudarshan E 2019 Smart healthcare monitoring system using raspberry Pi on IoT platform ARPN Journal of Engineering and Applied Sciences 14(4) 872-876.