SN Computer Science (2020) 1:175
https://doi.org/10.1007/s42979-020-00180-5

ORIGINAL RESEARCH

A Comparative Study and Analysis of Time Series Forecasting Techniques

Srihari Athiyarath[1] · **Mousumi Paul**[1] · **Srivatsa Krishnaswamy**[1]

Received: 19 April 2020 / Accepted: 21 April 2020 / Published online: 16 May 2020

Abstract
Time series data abound in many realistic domains. The proper study and analysis of time series data help to make important decisions. Study of such data is very useful in many applications where there are trendy changes with time or specific seasonality as in electricity demand, cloud workload, weather and sales, cost of business products, etc. By understanding the nature of the time series and the objective of analysis, we have used different approaches to learn and extract meaningful information that can satisfy the business needs. The present paper covers and compares various forecasting algorithmic approaches and explores their limitations and usefulness for different types of time series data in different domains.

Keywords Time series forecast · Deep learning · ARIMA · MVFTS · CNN · LSTM · CBLSTM

Introduction

The concept of time series analysis has come a long way right since its inception. Plenty of research has been done in time series analysis to accomplish multiple objectives. Even in today's machine learning and deep learning era, time series forecasting plays a crucial role to make important business decisions. Since the time series data are in real time, it can be used in organizations such as power management, stock market, health care, business, marketing, weather forecasting and many more. Time series prediction is essentially a part of temporal data mining and statistics. It is the process of careful collection and rigorous study of data that has been collected over a continuous period and development of a proper model that describes the inherent trend of the series. Time series forecasting is a process to predict the future with the help of history data. The assumption of time series forecasting is that the information will repeat itself in near future. The core intent of this paper is to synthesize and survey all crucial categories of time series methods that have been developed in areas of applied mathematics, technology, operations research and economic sciences. The objective of the work that follows is to create awareness of the time series analysis in various domains using the existing relevant work already available. By doing so, we hope to alleviate the communication gap between different disciplines considering how difficult this task is. The other motives of the paper are to explore the utility of different approaches of time series forecasting. Section 2 will explain the high-level approach to analyze the time series data. Section 3 will describe the different time series methods classified into the five categories. In Sect. 4, results and discussion have been provided for three different types of time series forecasting on different datasets (short-period forecasting, mid-period forecasting and long-period forecasting). Section 5 will conclude this paper.

Time Series Analysis

Time series data basically consist of systematic and non-systematic components. A systematic component has a consistency and recurrence that can be described by the model.

This article is part of the topical collection "Advances in Computational Approaches for Artificial Intelligence, Image Processing, IoT and Cloud Applications" guest edited by Bhanu Prakash K N and M. Shivakumar.

✉ Mousumi Paul
mousumipaul.88@gmail.com

Srihari Athiyarath
srihari.athiyarath@gmail.com

Srivatsa Krishnaswamy
sk@vigyanlabs.com

1 Vigyanlabs Innovation Private Limited, Mysore, Karnataka, India

A non-systematic component is a random phenomenon that cannot be directly modeled. The systematic component can be thought of a combination of level, trend, seasonality and the non-systematic component as a random variable or noise. Therefore, a decomposition technique extracts the trend and seasonal factors from time series data that helps to build the forecasting model. The more extensive decomposition includes local change effects, long-run effects, seasonal effects, etc. The removal of seasonal effects from data makes clear visibility of trend. The decomposition includes ways to find level, trend, seasonality, and noise from the data. The level is considered as mean pattern of the series while trend is harmonic pattern of the series. Seasonality is a very interesting characteristic, it is a short-term cyclic change over a long time series data, and finally some uncertain variation represents "noise" in time series data. All these four components may exist in additive or multiplicative form in the series data. However, the ultimate objective is to extract various patterns and an analysis can be made for depicting forecasts or future predictions which in turn helps in the growth of the business. But in this survey, prediction concept and the algorithms are the main foci. The kind of data that is considered here is called temporal data as it changes frequently with time and is represented using time stamps.

Time Series Forecasting Techniques

Various methods have been applied for time series forecasting. We have divided those into five categories.

Regression Method

One of the most widely used and easy to implement methods is the regression methods [1] and is used when a relationship is drawn between the predicted value and other factors. A presumption is made a way that the data to be predicted are divided into linear data and standard base data dependent on the factors that are affecting the data. For example, in electrical load forecasting problem, the load depends on the following factors such as the customer characteristics, the weekdays, the months and the conditions of weather. A design matrix $\left(X_t\right)$ is built in linear model which includes current and past observations of predictors order by time (t). The multiple linear regression model (MLR) is used $y(t) = x(t)b + u(t)$ to get an estimation of a linear relationship of the response $y\ (t)$. Here β acts as linear parameter, and $u\ (t)$ represents innovation term. The $x\ (t)$ matrix is defined based on the factors on which time series data change, for example local time, weekdays, months, weather, holiday information, etc. The error term is used to account for the variability in y that cannot be explained by the linear relationship between x and y.

Stochastic Approach

Due to the sequential nature, stochastic approach is widely used to draw inferences about the characteristics of the time series data. Such inference is subjected to certain laws of probability that can be described in terms of random variables $\left(x_1, x_2, \ldots, x_n\right)$, where each x corresponds to one instant of time. Assuming joint normality, then there are only n observations but $n + (n+1)n/2$ unknown parameters. To infer such a general probability structure from just one realization of the stochastic process will be impossible. Hence, stationary assumption is used to simplify the probability structure. A time series in which successive values are autocorrelated can be represented as a linear combination (linear filter) of a sequence of uncorrelated random variables. The linear filter representation is given by

$$z_t - \mu = x_t + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots = \sum_{j=0}^{\infty} \varphi_j x_{t-j}, \quad \varphi_0 = 1 \tag{1}$$

The random variable $\left\{x_t; t = 0, \pm 1, \pm 2, \ldots\right\}$ is a sequence of uncorrelated random variable from a fixed distribution with mean

$$^6E(x_t) = 0, \quad \mathrm{Var}(x_t) = E(x_t^2) = \sigma^2 \quad \text{and}$$
$$\mathrm{Cov}(x_t, x_{t-k}) = E(x_k, x_{t-k}) = 0 \quad \forall k \neq 0.$$

Seasonality and trend are the important characteristics of time series data. ARIMA [2], ARMA, SARIMA, etc., are one of such popular time series forecasting models. ARIMA [3–5] stands for autoregressive integrated moving average which has three parameters, namely (p, d, q). The model can be separated into three parts, autoregression (AR), which determines a relationship between the number of lagged observation and the observation itself; the number of lag observations included in the model is one of the main parameters (p). Integrated (I) which calculates the difference of raw observation used to make the time stationary and the degree of differencing is denoted by (d), and finally, the moving average (MA) uses the dependency between an observation and a residual error from a moving average model where the (MA) window is (q). SARIMA [6–8] is a variation of ARIMA which visualizes as seasonal autoregressive integrated moving average model. SARIMA is basically defined as (p, d, q) × (P, D, Q) where (p, d, q) and (P, D, Q) bear the same meaning of ARIMA and are divided into nonseasonal and seasonal parts, respectively. This model can also be represented as

$$\Phi(B^s)\varphi(B)(x_t - \mu) = \Theta(B^s)\theta(B)w_t \quad (2)$$

Here the nonseasonal autoregressive and moving average parts are $\varphi(B) = 1 - \varphi B - \cdots - \varphi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \cdots \theta_q B^q$, respectively, while on the other hand seasonal autoregressive and moving average parts are $\Phi(B^s) = 1 - \Phi_1 B^s - \cdots \Phi_p B^{PS}$ and $\Theta(B^s) = 1 + \Theta_1 B^s + \cdots + \Theta_Q B^{QS}$, respectively.

Soft Computing Techniques

In this deep learning era, RNN and CNN are the two widely used soft computing approaches in forecasting. We see that from the paper [9] RNN (LSTM) models are one of the best models for extraction patterns of the input features and used to span over a long sequence. They can model problems seamlessly with multiple input variables and this adds as a great benefit in time series forecasting, where simple classic regression methods can be difficult to adapt to multiple input and multivariate forecasting problems. In simple words, LSTM provides a lot of flexibility over other methods. LSTMs are a form of recurrent neural networks (RNNs). They have the ability to retain information for long time with the help of their inner cells which can carry information unchanged. The network has control of the cell state where it can edit, add or delete information using gates as shown below. Mathematically, the main aim of LSTM as shown in Eq. (3) is to find the conditional probability $p\ (y_1,\ldots, y_N|\ x_1,\ldots, x_N)$ where $(x_1,\ldots, x_N)$ are the input sequences and $(y_1,\ldots, y_N)$ are the corresponding output sequences in fixed-dimensional representation. In case of LSTM, the model is learning one part of the time series prediction and wants to apply it to another part which is not able to find local changes in time series. Hence, we can conclude that LSTMs are not supposedly good for short-period forecasting where learning on linearities and their "stationarity" is less brittle.

$$p(y_1, \ldots, y_N | x_1, \ldots, x_N) = \prod_{t=1}^{N} p(y_t | v, y_1, \ldots, y_t - 1) \quad (3)$$

On the other hand, CNN is very useful to capture local trends. From [10], we observe that CNNs that are mainly used for object detection were able to perform well enough when it came to time series analysis as well as it was able to mine and generate deep image features. We first performed pooling and convolution operations alternatively on the raw data and then connected it to the multilevel perceptron (MLP) for classification. The CNN architecture consists of a five-core layer such as input layer, convolution layer, neural layer, dense layer and pooling layer. In practically, it is found that CNN has performed well to adapt recent trends in time series data. In 2017, Rui Zhao and Ruqiang Yan implemented a combination of CNN and LSTM called as CBLSTM [11] that outperformed over all the existing techniques. In CBLSTMs, CNN acts as local feature extractor and bidirectional LSTM as temporal encoder. The model structure is shown in Fig. 1.

The performance of CBLSTM is pretty promising in case of long-term as well as short-term forecasting.

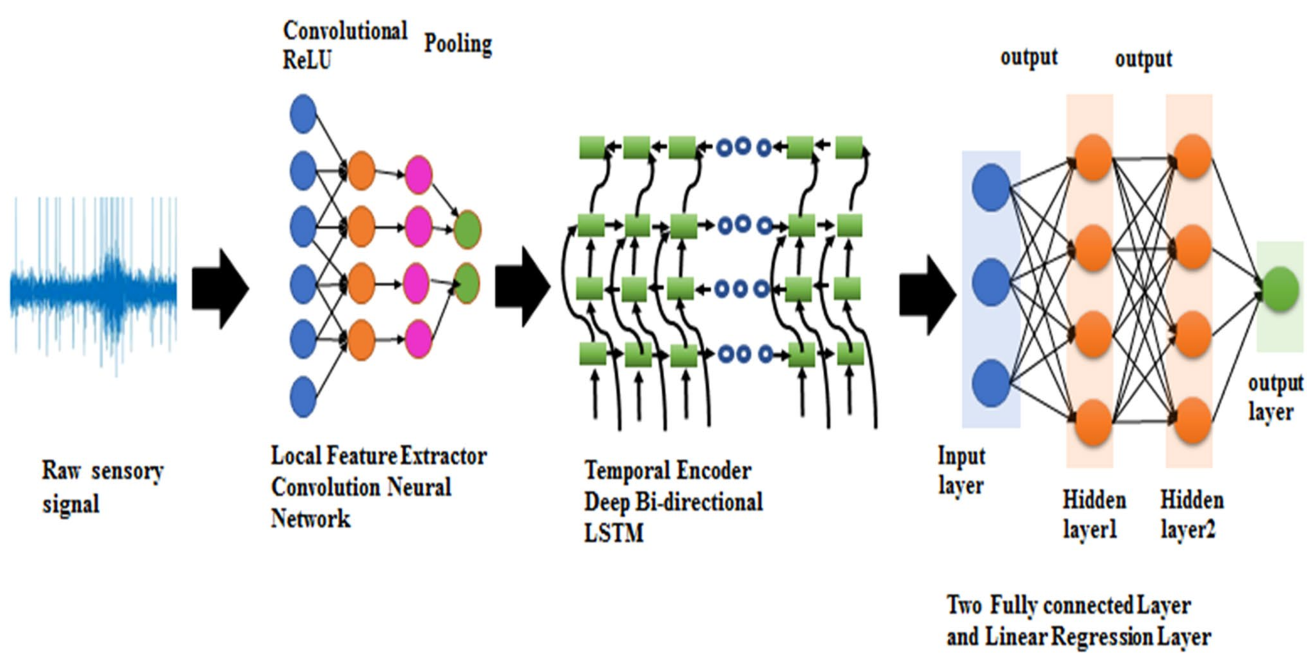

Fig. 1 Framework of convolutional bidirectional long short-term memory networks (CBLSTM) [11]

Fuzzy Logic Forecasting

Time series data are basically a one-dimensional array, but in many practical fields, it is an outcome of multiple variables. This type of data is called multivariate time series data. For example, stock exchange data [2] which have various attributes in the same day are a good example of such data. Conventional time series analysis cannot be used for this kind of data. A much more fitting way to design this type of time series would be to use the concept of fuzzy logic [12, 13]. There are three major stages in the analysis of the fuzzy logic-based time series forecasting. These include the fuzzification phase that defines the fuzzy rules and regulations, and de-fuzzification phase. Many novel methods have also been proposed. In [13], the author combined F transform and perceptron-based logical deduction to analyze and forecast the time data that has removed the prolongation of a prespecified curve. In [14], a weighted fuzzy model has been applied on Taiwan stock index data that proved weighted model performed better than conventional fuzzy time series models. In [15, 16], particle swarm optimization has been introduced with fuzzy *NN* to optimally find the intervals which increases the forecasting accuracy.

Results and Discussions

For evaluation purpose, we have performed three different types of time series forecasting, namely short-period forecasting, medium-period forecasting and long-period forecasting. Short-period forecasting refers to a time frame that ranges from a few days to a couple of weeks, medium period refers to a few months, and the long period considers a time period of more than a couple of years.

Datasets

We have used three discrete datasets, namely:

1. *Appliances energy prediction dataset* [17] This is a dataset consisting of a time and the energy consumption and observed that it had seasonality and stationarity but no trend.
2. *Solar flare dataset* [18] This dataset showed trend and mediocre stationarity, but no seasonality and a constant rise or drop was observed.
3. *Online retail dataset* [18] This dataset when plotted showed both seasonality but no trend or seasonality.

We have applied regression, LSTM, ARIMA, CNN, fuzzy-based method weighted MVFTS and CBLSTM on all these three discrete datasets and observed the following.

The performance of the above-mentioned methods for short-period forecasting, over all the three different datasets show that the best results by LSTMs, CNNs and weighted MVFTS and CBLSTM. ARIMA too performed significantly well; however, regression was not able to perform well. It is clear from the high RMSE and R2 values that regression is the most unlikely method for short-period forecasting (Figs. 2, 3).

In the case of medium-period forecasting, it is clear from Tables 1, 2, 3 well as the graphs that regression continues to fail and perform poorly where are ARIMA and weight MVFTS and CBLSTM performed extraordinarily well and likewise CNN and LSTMs do moderately well (Figs. 4, 5).

Finally, in long-period forecasting, over all the three different datasets show that the best results were similar to medium period, ARIMA, weighted MVFTS, CBLSTM and CNN consistently performed well. From the very high RMSE and R2 values, it is evident that regression is the most ineffective method to use (Figs. 6, 7).

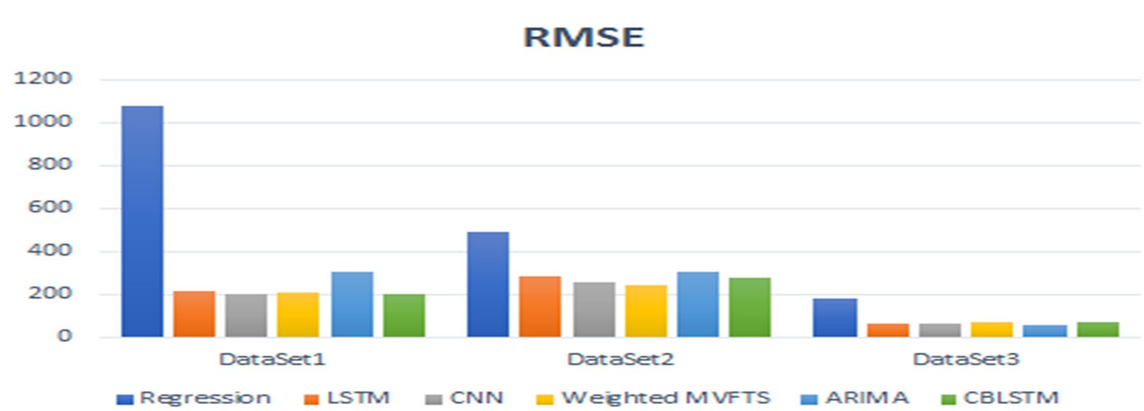

Fig. 2 RMSE values of the five different methods used for short-period forecasting

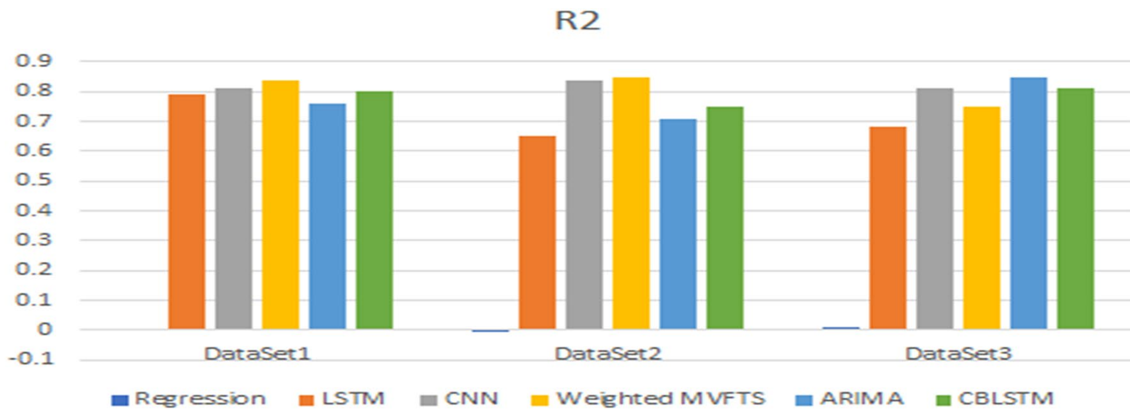

Fig. 3 Coefficient of determination (*R*2) values of the five different methods used for short-term forecasting

Table 1 Performance of short-term time series forecasting

Method name	Dataset 1		Dataset 2		Dataset 3	
	RMSE	R2	RMSE	R2	RMSE	R2
Regression	1080	0.001	489	−0.0023	182	0.012
LSTM	212	0.79	286	0.65	65	0.68
CNN	204	0.812	256	0.84	63	0.81
Weighted MVFTS	208	0.84	243	0.85	69	0.75
ARIMA	302	0.76	302	0.71	55	0.85
CBLSTM	201	0.8	274	0.75	68	0.81

Table 2 Performance of midterm time series forecasting

Method name	Dataset 1		Dataset 2		Dataset 3	
	RMSE	R2	RMSE	R2	RMSE	R2
Regression	947	0.0012	450	0.15	231	0.23
LSTM	312	0.34	245	0.4	185	0.69
CNN	302	0.82	225	0.81	142	0.89
Weighted MVFTS	150	0.81	214	0.83	135	0.84
ARIMA	165	0.74	203	0.84	136	0.86
CBLSTM	168	0.75	200	0.85	132	0.87

Table 3 Performance of long-term time series forecasting

Method name	Dataset 1		Dataset 2		Dataset 3	
	RMSE	R2	RMSE	R2	RMSE	R2
Regression	1161	0.00014	291	0.015	310	0.32
LSTM	615	0.31	186	0.36	243	0.51
CNN	312	0.75	145	0.71	200	0.68
Weighted MVFTS	300	0.79	164	0.68	212	0.75
ARIMA	289	0.84	135	0.81	185	0.84
CBLSTM	256	0.91	142	0.79	175	0.92

Conclusion

In today's world, there are a wide variety of methods that can be utilized to fit forecasting, control methods and estimations which involve time series. In this paper, we have described, implemented and summarized different forecasting models. We have analyzed and predicted using different time series

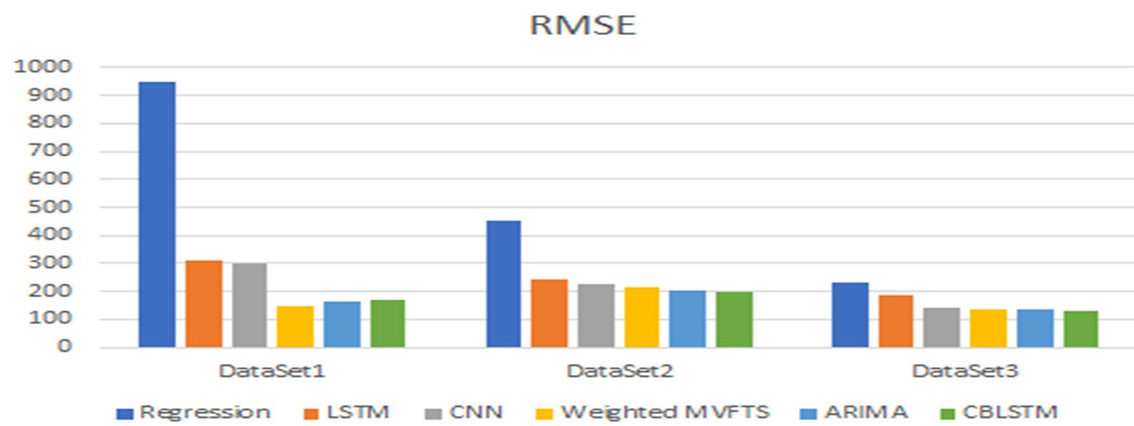

Fig. 4 RMSE values of the five different methods used for midterm forecasting

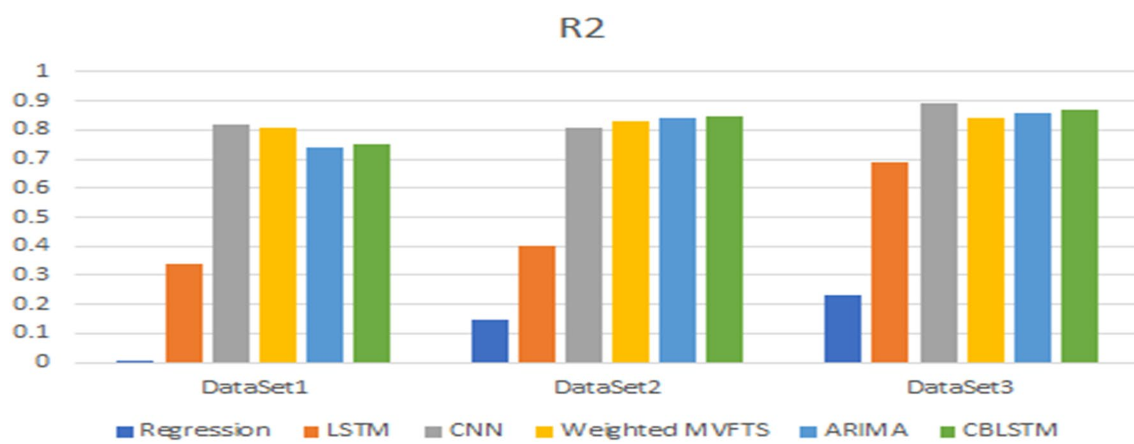

Fig. 5 Coefficient of determination ($R2$) values of the five different methods used for midterm forecasting

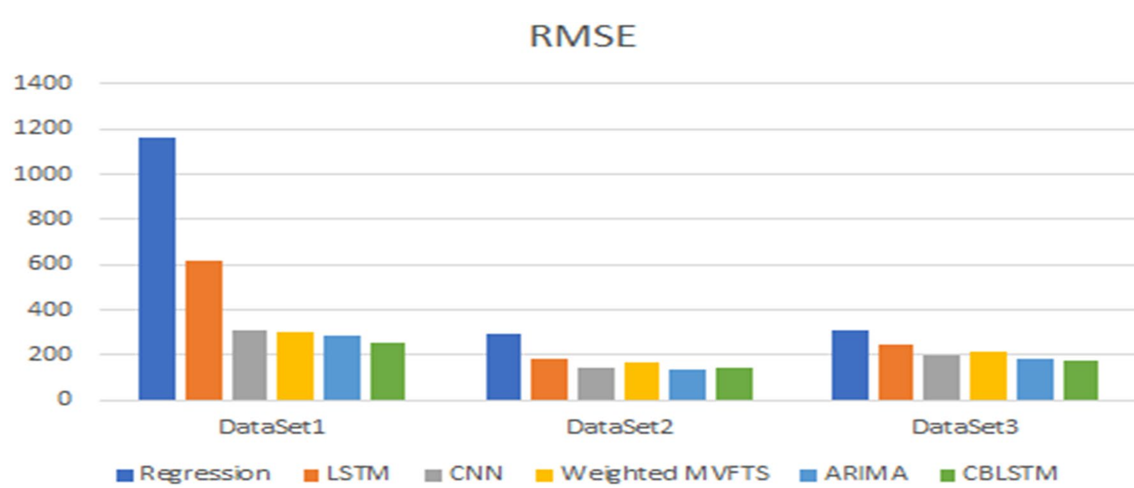

Fig. 6 RMSE values of the five different methods used for long-term forecasting

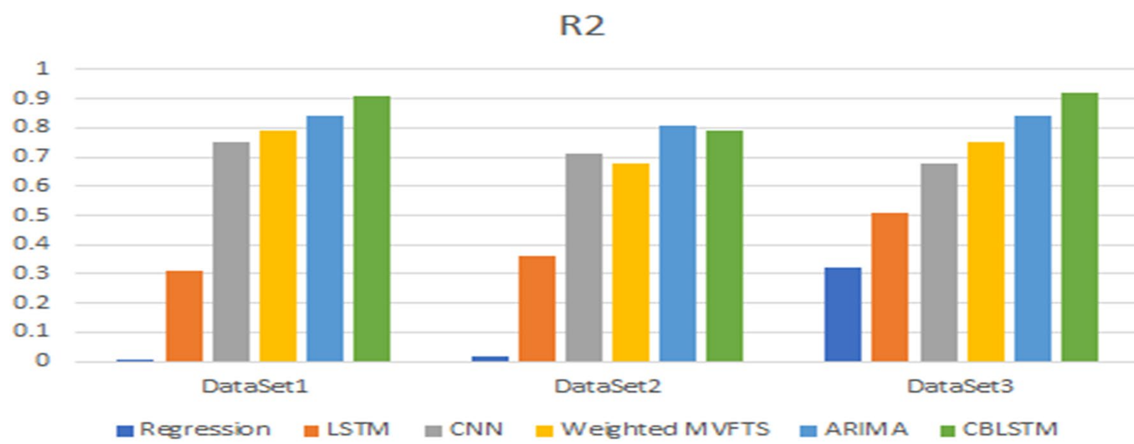

Fig. 7 Coefficient of determination ($R2$) values of the five different methods used for long-term forecasting

model based on a variety of discrete datasets considering its stationarity, seasonality as well as its trend. As per our survey, it is proven that the performance of the models depends on the time space (short, mid or long). For short-range LSTMs, CNN and weighted MVFTS, CBLSTM performed quite well, whereas for midterm as well as long term, ARIMA and MVFTS, CBLSTM provided momentous results. From a satisfactory understanding about the considered forecasting models, we can conclude that linear regression failed to perform in all the time spaces while weighted MVFTS and CBLSTM outperformed the rest of the models. Our future work will be to develop a hybrid model and carry out an empirical study to identify the top forecasting techniques in the field of power management and green computing.

Funding This work was supported by the Vigyanlabs Innovations Private Limited.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Seber GAF, Lee AJ. Linear regression analysis. Wiley series in probability and statistics. New York: Wiley; 2012.
2. Song Q, Leland RP, Chissom BS. A new fuzzy time-series model of fuzzy number observations. Fuzzy Sets Syst. 1995;73(3):341–8.
3. Box GEP, Jenkins G. Time series analysis, forecasting and control. San Francisco: Holden-Day; 1970.
4. Cochrane JH. Time series for macroeconomics and finance. Chicago: Graduate School of Business, University of Chicago; 1997.
5. Hipel KW, McLeod AI. Time series modelling of water resources and environmental systems. Amsterdam: Elsevier; 1994.
6. Vagropoulos SI, Chouliaras GI, Kardakos EG, Simoglou CK, Bakirtzis AG. Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. In: IEEE international energy conference (ENERGYCON), Leuven 1–6; 2016.
7. Hamzacebi C. Improving artificial neural networks' performance in seasonal time series forecasting. Inf Sci. 2008;178:4550–9.
8. Kihoro JM, Otieno RO, Wafula C. Seasonal time series forecasting: a comparative study of ARIMA and ANN models. Afr J Sci Technol Sci Eng Ser. 2004;5(2):41–9.
9. Faustryjak D, Jackowska-Strumiłło L, Majchrowicz M. Forward forecast of stock prices using LSTM neural networks with statistical analysis of published messages. In: Interdisciplinary PhD workshop (IIPhDW) 2018 international, pp 288–292; 2018.
10. Yang J, Nguyen MN, San PP, Li XL. Deep convolutional neural networks on multichannel time series for human activity recognition. In: Fourth international joint, 2015—aaai.org.
11. Zhao R, Yan R, Wang J, Mao K. Learning to monitor machine health with convolutional bi-directional LSTM networks. Sensors. 2017;17(2):273.
12. Cheng C-H, Cheng G-W, Wang J-W. Multi-attribute fuzzy time series method based on fuzzy clustering. Expert Syst Appl. 2008;34(2):1235–42.
13. Wang J-W, Liu J-W. Weighted fuzzy time series forecasting model. Berlin: Springer; 2010. p. 408–15. ISBN 978-3-642-12145-6.
14. Yu T. Weighted fuzzy time series models for TAIEX forecasting. Physica A. 2005;349:609–24. https://doi.org/10.1016/j.physa.2004.11.006.
15. Bose M, Mali K. Fuzzy time series forecasting model using particle swarm optimization and neural network. In: SocProS 2017, vol 1. https://doi.org/10.1007/978-981-13-1592-3_32.
16. Qiu W, Zhang C, Ping Z. Generalized fuzzy time series forecasting model enhanced with particle swarm optimization. Int J U E-Serv Sci Technol. 2015;8:129–40. https://doi.org/10.14257/ijunesst.2015.8.5.11.
17. Candanedo LM, Feldheim V, Deramaix D. Data driven prediction models of energy use of appliances in a low-energy house. Energy Build. 2017;140:81–97.
18. Dua D, Graff C. UCI machine learning repository http://archive.ics.uci.edu/ml. Irvine, CA: University of California, School of Information and Computer Science. http://archive.ics.uci.edu/ml; 2019.