JEAS ,Vol. 31 No. 148 (2025) pp (171-181)
P-ISSN 2518-5764 - E-ISSN 2227-703X
DOI: https://doi.org/10.33095/j87hvg90

Journal of Economics and Administrative Sciences

Using Convolutional Neural Networks to Estimate Missing Values in Univariate Time Series Data

Wasn Saad Mahdi
College of Political Science; University of Baghdad; Baghdad, Iraq.

Received:23/4/2025 **Accepted:9/7/2025** **Published: 1/8/2025**

Abstract:
In this research, deep neural networks (CNNs) were used to estimate missing values in a univariate time series dataset and compared with DES/Holt and (SVR) estimation methods to determine their accuracy in handling missing values in the dataset. These missing values directly affect the process of building mathematical and statistical models, affecting the accuracy of the final results that enable us to make the right decisions in the future. The presence of missing values in the dataset is due to problems that occur during sampling, such as a malfunction in the measuring device, security attacks, or communication errors. The results indicated that the CNN model outperformed other methods by using a simulation approach that generated data randomly. Three different sample sizes (60, 100, 300) were selected with missing data percentages of (10%, 15%, 20%) for each, meeting the "missing at random" (MAR) condition. The Box-Jenkins MA (1) model was applied once with $\beta = 0.5$ and again with $\beta = 0.9$. The accuracy of the estimation methods was evaluated using the accuracy criteria mean sum of squared error (MSE) and root mean squared error (RMSE).
Keywords: Convolutional Neural Networks, univariate time series, Missing value, Deep learning.

1. **Introduction:**

Research in all fields requires a complete dataset (i.e., no missing values). However, researchers often face a critical problem: incomplete data (i.e., missing values), which is the main reason for not providing accurate results and reaching a final decision. There are three types of missing values in a dataset: missing not at random (MNAR), missing at random (MAR), and missing completely at random (MCAR). In this research, we will examine a dataset containing missing values from a univariate time series. This is the main reason for distorting the statistical properties of time series data, and mishandling these values can lead to inaccurate data analysis results. Appropriate methods must be chosen to handle missing values, whether using traditional statistical methods or artificial intelligence techniques, to obtain more reliable and accurate time series analysis results. This research uses a traditional statistical method, Holt's method, and two machine learning techniques (Ahn, Sun and Kim, 2022). Machine learning, is a branch of artificial intelligence concerned with creating algorithms that enable computers to learn from (training data). We can also define it as the process of solving a problem by building an algorithmic statistical model based on the data set that is collected, and this model is used to solve the problem (Burkov, 2019).

2. **Literature Review and Hypothesis Development:**

Time series data contains two types of values: outliers and missing values. These present fundamental challenges in data analysis, as these values can significantly reduce the consistency of data collected from devices due to non-response or error, thus reducing the accuracy of predictive models and statistical processes. Missing values in univariate time series pose a significant challenge in terms of their difficulty in processing, as each value depends on the accuracy of the time series, as it represents the evolution of the phenomenon over time. Over the years, numerous studies have proposed various methods for processing missing values, including traditional or classical statistical methods and machine learning methods. Studies indicate that machine learning methods demonstrate superior performance. We will review some of these studies.(Al-Milli and Almobaideen, 2019) They proposed a hybrid deep learning neural network (Jordan network) that was used as a missing data prediction model with a genetic algorithm to optimize the neural network weights to impute missing data for medical IoT applications. (Kim, Ko and Kim, 2019)Theses researchers used four methods the machine learning and statistical methods to estimate missing data in weather data and found that the k-nearest neighbors method is the most appropriate method for estimating missing data.(Chaudhry et al., 2019) The researchers proposed a new method that transforms a single variable into a multivariate model by exploiting the high seasonality and random absence of this variable. They compared the proposed method with Kalman filtering and the default multivariate imputation method. Performance evaluation results clearly show that the proposed method significantly outperforms Kalman filtering and the default method in terms of imputation and prediction accuracy. It overcomes the challenges of handling missing values in univariate and multivariate time series datasets. (Flores, Tito and Silva, 2019) The researchers used two simple algorithms to estimate missing value in univariate time series, which are based on the means of the nearest neighbors (Local Average of Neighbors Neighbors(LANN), Local Average of Neighbors Neighbors+(LANN+)) . (Caillault, Lefebvre and Bigand, 2020)Propose a method for estimating large missing intervals using a dynamic time wrapping (DTW) algorithm that searches for values most similar to values preceding or following missing values in univariate time series data with autocorrelation, high correlation, strong seasonality, and complex distribution. The researcher (Ding *et al.*, 2020) used three interpolation algorithms, including (1) radial basis functions, (2) moving least squares (MLS), and (3) weighted adaptive inverse distance to estimate missing values in eight selected IoT time series datasets, and in their experiment, they indicated that MLS-based estimation is the best. (Phan, 2020)Researchers proposed a new method for handling consecutive missing values in univariate time series using machine learning (ML) techniques, called MLBUI.

First, for each missing value, they transformed the data into a multivariate time series using data before and after the missing value. After this transformation, machine learning techniques were used to estimate the missing values, applying them to forward and backward prediction, and then processing the missing values using the median values of the two prediction sets. They found that MLBUI outperformed several existing methods. The researchers (Contractor and Roughan, 2021) used a novel contextual neural network with long short-term memory (LSTM) to estimate missing values in the temperature data series, using the pre- and post-gap data as separate feature variables. (Saeipourdizaj, Sarbakhsh and Gholampour, 2021) The researchers used multiple methods such as Methods of mean، EM algorithm، Moving average, K-nearest neighbor (KNN), Predictive mean matching (PMM), Interpolation, regression, classification, and tree regression to estimate the missing values of the data on pollutants in the air, PM10 and O3. the Interpolation method and Moving average and K-nearest neighbor (KNN) performed best because these methods depend on the next and previous information.(Savarimuthu and Karesiddaiah, 2021) The researchers suggested an iterative algorithm based on clustering univariate time series data, taking into account the characteristics of the general trend, seasonality, cyclicity, and residuals. The method relies on the nearest neighbor method, which is based on the similarity principle within each group, to handle missing values.(Yang et al., 2022) To estimate missing values in time series of taxi traffic data, researchers presented a generative adversarial network (ST-FVGAN), where the model is built from a generative network that relies on spatial and temporal correlation and external factors. They found that this network (ST-FVGAN) has high accuracy in estimating missing values.(Ahn, Sun and Kim, 2022) Researchers have used multiple methods to handle missing values and the results show that the k-nearest neighbor method is the most efficient at handling missing values, compared to other imputation methods.(Chhabra, 2023) Use methods such as (mean, last observation carried forward , Kalman smoothing , seasonal decomposition using interpolation , seasonal imputation using mean , moving average and moving average with exponential weighting , linear interpolation , spline interpolation and stine interpolation to address missing values in univariate time series data with seasonal characteristics. It was found that the methods “seasonal decomposition using interpolation and linear interpolation and Kalman smoothing" gave a better performance than other methods on time series with seasonal charactering. (Niako et al., 2024)The researchers used methods such as averaging, Kalman filtering, linear integrals, curve integrals, Steinman integrals, exponential weighted moving average (EWMA), simple moving average (SMA), k-nearest neighbor (KNN), and last observation carried forward (LOCF) inference techniques on the time series structure, and evaluated the prediction performance of LSTM and ARIMA models. The results showed that the ARIMA model performs better on data with higher autocorrelation.

3. Research Methodology:

3.1 (DES / Holt) Double exponential smoothing / Holt

Holt proposed this method in 1957. It relies on two parameters (a trend parameter and a level parameter). This method is used to estimate missing values in univariate time series. For more details, see here.(Mahdi and ALmohana, 2022) (Rachmat and Suhartono, 2020) (Cadenas, Jaramillo and Rivera, 2010).

- The trend parameter can be calculated according to the equation below:

$H_t = \beta(L_t - L_{t-1}) + (1-\beta)H_{t-1}$ … (1)

Where: H_t : trend parameter in the time period t.

H_{t-1} : trend parameter in the previous time period t-1.

β : exponential smoothing constant has a value between 0< β<1.

L_t: level parameter in the time period t.

L_{t-1}: level parameter in the previous time period t-1.

- The level parameter can be calculated according to the equation below:

$$L_t = \alpha w_t + (1-\alpha)(L_{t-1} + H_{t-1}) \quad \dots (2)$$

Where: α :exponential smoothing coefficient and its value is between 0< α<1.
w_t: time series variable.

- The basic equation of the Holt method is as follows:

$$\widehat{W}_{t+c} = L_t + H_t(C) \quad c = 1,2,\dots,n \quad \dots \ (3)$$

Where: $\widehat{W}_{t+c}$:estimate of the missing value in the subsequent $t+c$ time intervals.
C: number of time periods to estimate in the future.11

Algorithm steps (DES / Holt):

1- Define the value β.
2- Define the value α.
3- define the time series variable w_t.
4- Specify a value n.
5- Specify a value C.
6- Enter the equation The trend parameter، Equation (1).
7- Enter the equation The level parameter، Equation (2).
8- Enter the equation The basic equation of the Holt method، Equation (3).
9- Stop.
10- Output.

3.2 Convolutional neural networks (CNN) :

Deep Learning (DL) is a type of machine learning used to train a deep neural network. It is called deep because it contains two or more hidden layers. Bates was the first to create deep neural networks in 1943, but he struggled to train them effectively. However, in 1984, researcher Geoffrey Hinton was the first to successfully train a deep neural network. In recent years, deep learning has gained significant importance, having successfully solved some of the problems facing artificial intelligence (Heaton, 2015a) **.** Deep neural networks mimic human thinking, making them capable of gaining a lot of knowledge from a vast and massive dataset.(Adlia and Mahmoud, 2024). Some of the deep learning methods chosen to address missing values in time series datasets such as CNN are the first convolutional network to be successfully deployed since the advent of deep learning. Its work is derived from the biological processes present within the visual section of the human brain. It was called convolution due to its basic component of convolutional layers that use mathematical functions for convolution. It has proven its effectiveness in many fields such as analysis (images, audio, text, and time series). CNN are used to address the problem of missing values in univariate time series. They use a function to plot a series of past data and use it as input to process the output. In this research, we use a one-dimensional convolutional neural network (CNN-1D) because the data in time series is univariate, meaning that there is data for one variable and a number of time steps. The CNN consists of the following of three hidden neural layers (Shaikh, Rasool and Lone, 2022) : Convolutional Neural Network Architecture

- Input layer:

In this layer, input Univariate time series data that contains missing values .

$$X = x_1, x_2, \dots, x_n \ \dots (4)$$

- Convolution-1D layer (Heaton, 2015b):

Usually, a single convolutional hidden layer is used, but in some cases where the time series is very long, a second convolutional hidden layer may be used. This layer convolves the input data and transmits its output to the next layer, the pooling layer. Equation of this layer is(Chatterjee *et al.*, 2025):

$$Y_t^i = \sum_{k=0}^{k-1} W_t[k] \,.\, X[i+k] + b_t$$

Where Y_t^i : **is the value output of the feature map at position i.**
K: size of filter.
$W_t[k]$: is the weight of the filter at position k.
b_t: is the bias term of the filter.

- Activation function we use (Relu) It is an activation function that we use a lot in deep learning algorithms because it helps in reducing the problem of vanishing gradient because it deletes all negative values and gives very fast results. Its value range is [0, x) and its output is zero when the input is negative, but if the input is greater than or equal to zero, then it is the same value. The ReLU applies the function(Ren *et al.*, 2018):

$R_t^i = \max[0, Y_t^i)$… (5)

- Maxpooling layer:

this layer reduce computational complexity and spatial dimensions(Sahu, 2023).These are layers located after the convolutional layer, and their function is to reduce dimensions and computational cost (Cerqueira and Roque, 2024)(Chatterjee *et al.*, 2025).

$P_t^i = \max(Y_t^i, Y_{t+1}^i, \dots, Y_{t+p-1}^i)$ … (6)

- output layer

▪ fully connected or linear to output node (output of the convolutional layer), It collects information from previous layers to make predictions.

▪ Flatten Layers: The results obtained from the convolution and pooling layers are flattened and sent to a dense layer.

- Dense Layers

This is the final layer on neural networks We determine through it Neuron Count and Activation Function such as (ReLU, Sigmoid or Hyperbolic tangent)

Algorithm steps (CNN):

1- Input $X = x_1, x_2, \dots, x_n$.
2- Convolution-1D layer Specify (filters = 32, kernel size= 5, activation =ReLU).
3- Specify Maxpooling layer or pool size = 2.
4- Specify a Flatten.
5- Specify a Dense = 1
6- Output.

3.3 Support Vector Regression (SVR)

It is one of the supervised machine learning algorithms used for regression and classification models, and for finding outliers. It is used to estimate missing values in univariate time series datasets. Each point is represented by a point with multiple input values and a single output value. This method creates a dividing line (called a hyperplane) denoted by $\hat{Z}0$ between the dataset to divide the observations into two groups equally distributed around the values to be estimated. Each group is independent of the other, provided that the dataset observations do not exceed the two lines (Z1, Z2). Any observations outside the two lines are ignored, thus eliminating any errors (Gazzola and Jeong, 2021)(Honghai *et al.*, 2005)(Li *et al.*, 2009). The goal of this method is to maximize the distance between the two lines ($y1$, $y2$).

- The equation of the dividing line is:

$\hat{Z} = b + S^T\text{y} \ldots (7)$

Where: $\hat{Z}$: The output is a real number to represent the function that estimates the missing value.
b: is a real number to represent the bias.
S^T: weight vector.
y : vector the input values.

Algorithm steps (SVR):

1- Input vector values y.
2- We divide the data set into two equal groups: $y1$ the first is a group of observations that do not contain missing values, and $y2$ the second is a group of observations that contain missing values.
3- choose $y1$ for the training set.
4- choose $y2$ to predict missing values.
5- Use of the equation number (4).
6- Stop.
7- Output.

Result:

To compare the methods used to estimate missing values:

- Python was used to build a simulation model and generate missing data at random (MAR). Simulation experiments were conducted using three sample sizes (100, 300, and 500) with repetition (r=500) for each experiment. The variable (X) was generated into numbers following a standard normal distribution, and random errors were generated following a standard normal distribution where the mean is zero and the variance is one. The simulation model is Box-Jenkins the MA (1) ($\hat{X}_t = e_t - \beta e_{t-1}$) is used once with a value of (β = 0.5) and with (β = 0.9). The missing data type is missing at random (MAR), meaning that the value of the missing observation depends on the time point of that observation but not on the value of the missing observation itself. It used Three missing percentages in the dataset: (10%, 15%, and 20%). To compare and determine the best estimation method, two accuracy measures are used, namely the mean square error (MSE) and the root mean square error (RMSE), with the smallest value of these measures being used.(Al-Mohana, Firas Ahmed, and Saleem, 2018)(Lim and Zohren, 2021).

4. Discussion of Results:

Table 1: MSE for the methods used to estimate missing values when the sample size is (100,300,500) with missing (10%,15%,20%), (β =0.5) and with repetition (r=500)

β =0.5					
	Missing	sample size	Methods		
			CNN	Holt	SVR
MSE	10%	100	1.26242	1.75452	1.36063
		300	1.15665	1.64944	1.35624
		500	1.07006	1.62362	1.41043
	15%	100	1.26139	1.73887	1.55446
		300	1.16576	1.6632	1.5599
		500	1.09349	1.65324	1.58311
	20%	100	1.27454	1.75501	1.7855
		300	1.17849	1.67425	1.76485
		500	1.1148	1.6722	1.73743

Source: Prepared by researcher based on Simulation.

Table 2: MSE for the methods used to estimate missing values when the sample size is (100,300,500) with missing (10%, 15%, 20%), (β =0.9) and with repetition (r=500)

β =0.9					
	Missing	sample size	Methods		
			CNN	Holt	SVR
MSE	10%	100	1.8162	2.40066	1.49975
		300	1.44907	2.27918	1.49117
		500	1.31863	2.23287	1.53265
	15%	100	1.83746	2.40244	1.74785
		300	1.63548	2.28523	1.71917
		500	1.38921	2.27272	1.7488
	20%	100	1.84805	2.43494	2.00142
		300	1.5749	2.2987	1.95722
		500	1.43891	2.29659	1.95341

Source: Prepared by researcher based on Simulation.

In tables (1), when the value of (β =0.5), we note the following:

i. The best method for estimating missing values in univariate time series is the deep convolutional neural network (CNN) method, which yielded the lowest mean sum of squared errors (MSE) for all data sizes (100, 300, 500), and all loss ratio values (10%, 15%, 20%).

ii. The SVR method, which follows the CNN method in its efficiency in estimating missing values, yielded the lowest mean sum of squared errors (MSE)

iii. However, the Holt method yielded the worst results, with MSE values among the highest compared to the other methods across all data sizes and loss ratios.

iv. The mean sum of squared errors in CNNs decreases with increasing data sizes and all loss ratio values.

v. The mean sum of squared errors (MSE) in SVR increases at sample sizes of 500 and loss ratios of (10% and 15%), while at loss ratios of (20%), it decreases at sizes (500) of (100 ,300).

vi. As for the loss ratio of 20%, we note that the MSE is highest at size (100) and decreases with increasing data size.

In tables (2), when the value of (β =0.9), we note the following:

i. The deep convolutional neural network (CNN) method is the best method for estimating missing values in univariate time series, yielding the lowest mean sum of squared errors (MSE) for all data sizes (100, 300, 500), and all loss ratio values (10%, 15%, 20%).

ii. The SVR method follows the CNN method in its efficiency in estimating missing values, as it also yielded the lowest mean sum of squared errors (MSE) compared to Holt's method.

iii. The mean sum of squared errors in SVR increases at sample sizes (500) and at loss ratios (10%, 15%), and decreases at sizes (300) and at the same loss ratio.

iv. The Holt method showed the worst results, with MSE values among the highest compared to the other methods at all sizes and for all loss ratios.

v. The mean sum of squared errors in SVR at sample sizes (100) and at loss ratios of 10% and 15% is lower in the method CNN.

Table 3: RMSE for the methods used to estimate the missing values when the sample size is (100,300,500) with missing (10%,15%,20%), (β =0.5) and with repetition (r=500)

β =0.5					
	Missing	sample size	Methods		
			CNN	Holt	SVR
RMSE	10%	100	1.12357	1.32458	1.16646
		300	1.07548	1.28431	1.16458
		500	1.03444	1.27421	1.18762
	15%	100	1.12312	1.31866	1.24678
		300	1.07970	1.28965	1.24896
		500	1.04570	1.28578	1.25822
	20%	100	1.12896	1.32477	1.33623
		300	1.08558	1.29393	1.32848
		500	1.05584	1.29314	1.31812

Source: Prepared by researcher based on Simulation.

In tables (3), when the value of (β =0.5), we note the following:

i. The best method for estimating missing values in univariate time series is the deep convolutional neural network (CNN) method, which yielded the lowest Root mean squared error (RMSE) for all data sizes (100, 300, 500), and all loss ratio values (10%, 15%, 20%).
ii. The SVR method, which follows the CNN method in its efficiency in estimating missing values, yielded the lowest (RMSE).
iii. However, the Holt method yielded the worst results, with MSE values among the highest compared to the other methods across all data sizes and loss ratios.
iv. The Root mean squared error in CNNs decreases with increasing data sizes and all loss ratio values.
v. The (RMSE) in SVR increases at sample sizes of 500 and loss ratios of (10% and 15%), while at loss ratios of (20%), it decreases at sizes (500) of (100 ,300).
vi. As for the loss ratio of 20%, we note that the RMSE is highest at size (100) and decreases with increasing data size in all methods.

Table 4: RMSE for the methods used to estimate the missing values when the sample size is (100,300,500) with missing (10%,15%,20%), (β =0.9) and with repetition (r=500)

β =0.9					
	Missing	sample size	Methods		
			CNN	Holt	SVR
RMSE	10%	100	1.34766	1.54941	1.22464
		300	1.20377	1.50970	1.22113
		500	1.14832	1.49428	1.23800
	15%	100	1.35553	1.54998	1.32206
		300	1.27886	1.51170	1.31117
		500	1.17865	1.50755	1.32242
	20%	100	1.35943	1.56043	1.41472
		300	1.25495	1.51615	1.39901
		500	1.19955	1.51545	1.39764

Source: Prepared by researcher based on Simulation.

In tables (4), when the value of (β =0.9), we note the following:

i. The deep convolutional neural network (CNN) method is the best method for estimating missing values in univariate time series, yielding the lowest Root mean squared error (RMSE) for all data sizes (100, 300, 500), and all loss ratio values (10%, 15%, 20%).
ii. The SVR method follows the CNN method in its efficiency in estimating missing values, as it also yielded the lowest (RMSE) compared to Holt's method.
iii. The Root mean squared error in SVR increases at sample sizes (500) and at loss ratios (10%, 15%), and decreases at sizes (300) and at the same loss ratio.
iv. The Root mean squared error in SVR at sample sizes (100) and at loss ratios of 10% and 15% is lower in the method CNN.
v. The Holt method showed the worst results, with RMSE values among the highest compared to the other methods at all sizes and for all loss ratios.

5. Conclusion:

The CNN method has proven it's the best in estimating missing values in univariate time series. This is demonstrated by the simulation results, which gave lower MSE values for all sample sizes, which had a clear impact on the accuracy of CNN results. The MSE and RMSE value decreased with increasing sample size at all loss ratios. It can also be noted that its performance is better with large sample sizes. The (SVR) method performed better with small sample sizes (100), giving the lowest sum of squared error value and Root mean squared error at the lowest loss ratios (10%, 15%).

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved by The Local Ethical Committee in The University.

References:

Ragheb Abdulrazak Adlia, A., & Hatim Mahmoud, A. (2024). The Role of Artificial Intelligence Techniques in Enhancing Project Completion Speed: A Study on Using LSTM Networks for Predicting Delay Times. *Journal of Economics and Administrative Sciences*, *30*(142), 197-219. https://doi.org/10.33095/n7rfbg91

Ahn, H., Sun, K. and Kim, K. P. (2022). Comparison of missing data imputation methods in time series forecasting. *Computers, Materials & Continua*, 70(1), pp. 767–779.

Al-Milli, N. and Almobaideen, W. (2019). Hybrid neural network to impute missing data for IoT applications. in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*. IEEE, pp. 121–125.

Al-Mohana, Firas Ahmed, and Saleem, N. (2018) Comparison Some Estimation Methods Of GM(1,1) Model With Missing Data and Practical Application. *Journal of Economics and Administrative Sciences*, *24*(103), 420. https://doi.org/10.33095/jeas.v24i103.121

Burkov, A. (2019). *The hundred-page machine learning book*. Quebec city, canada: Andriy Burkov Quebec City, QC, Canada.

Cadenas, E., Jaramillo, O. A. and Rivera, W. (2010). Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renewable Energy*, 35(5), pp. 925–930.

Caillault, É. P., Lefebvre, A. and Bigand, A. (2020). Dynamic time warping-based imputation for univariate time series data. *Pattern Recognition Letters*, 139, pp. 139–147.

Cerqueira, V. and Roque, L. (2024). *Deep Learning For Time Series Cookbook use PyTorch and Python recipes for forecasting, classification, and anomaly detection*.

Chaudhry, A. *et al.* (2019). A method for improving imputation and prediction accuracy of highly seasonal univariate data with large periods of missingness. *Wireless Communications and Mobile Computing*, 2019(1), p. 4039758.

Chhabra, G. (2023). Comparison of imputation methods for univariate time series. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11, pp. 286–292.

Contractor, S. and Roughan, M. (2021). Efficacy of feedforward and LSTM neural networks at predicting and gap filling coastal ocean timeseries: oxygen, nutrients, and temperature. *Frontiers in Marine Science*, 8, p. 637759.

Chatterjee, A. et al. (2025) 'Supervised Anomaly Detection in Univariate Time-Series Using 1D Convolutional Siamese Networks', *IEEE Access.*

Ding, Z. et al. (2020). Comparison of estimating missing values in iot time series data using different interpolation algorithms. *International Journal of Parallel Programming*, 48(3), pp. 534–548.

Flores, A., Tito, H. and Silva, C. (2019). Local average of nearest neighbors: Univariate time series imputation. *International Journal of Advanced Computer Science and Applications*, 10(8), pp. 45–50.

Gazzola, G. and Jeong, M. K. (2021). Support vector regression for polyhedral and missing data. *Annals of Operations Research*, 303(1–2), pp. 483–506. doi: 10.1007/s10479-020-03799-y.

Heaton, J. (2015a). AIFH, volume 3: deep learning and neural networks. *Journal of Chemical Information and Modeling*, 3.

Heaton, J. (2015b). Artificial Intelligence for Humans, Volume 3: Neural Networks and Deep Learning. sl: Heaton Research. *Inc.: Chesterfield, UK*.

Honghai, F. *et al*. (2005). A SVM regression based approach to filling in missing values, in *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Springer, pp. 581–587.

Kim, T., Ko, W. and Kim, J. (2019).Analysis and impact evaluation of missing data imputation in day-ahead PV generation forecasting. *Applied Sciences*, 9(1), p. 204.

Li, Q. *et al.* (2009). The investigation and application of SVC and SVR in handling missing values. in *2009 First International Conference on Information Science and Engineering*. IEEE, pp. 1002–1005.

Lim, B. and Zohren, S. (2021) 'Time-series forecasting with deep learning: a survey', *Philosophical Transactions of the Royal Society A*, 379(2194), p. 20200209.

Saad Mahdi, W., & A. Mohammed ALmohana, F. (2022). A Comparison of a Radial Basis Function Neural Network with other Methods for Estimating Missing Values in Univariate Time Series. *Journal of Economics and Administrative Sciences*, *28*(134), 134-146. https://doi.org/10.33095/jeas.v28i134.2426

Niako, N. *et al.* (2024). Effects of missing data imputation methods on univariate blood pressure time series data analysis and forecasting with ARIMA and LSTM. *BMC Medical Research Methodology*, 24(1), p. 320.

Phan, T.-T.-H. (2020). Machine learning for univariate time series imputation. in *2020 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*. IEEE, pp. 1–6.

Rachmat, R. and Suhartono, S. (2020). Comparative Analysis of Single Exponential Smoothing and Holt.s Method for Quality of Hospital Services Forecasting in General Hospital. *Bulletin of Computer Science and Electrical Engineering*, 1(2), pp. 80–86.

Ren, R. *et al.* (2018) 'Deep convolutional neural networks for log event classification on distributed cluster systems', in *2018 IEEE international conference on big data (big data)*. IEEE, pp. 1639–1646.

Saeipourdizaj, P., Sarbakhsh, P. and Gholampour, A. (2021). Application of imputation methods for missing values of pm10 and o3 data: Interpolation, moving average and k-nearest neighbor methods. *Environmental Health Engineering and Management*, 8(3), pp. 215–226. doi: 10.34172/EHEM.2021.25.

Sahu, N. (2023). Mathematics for Machine Learning_ A Deep Dive into Algorithms.

Savarimuthu, N. and Karesiddaiah, S. (2021). An unsupervised neural network approach for imputation of missing values in univariate time series data. *Concurrency and Computation: Practice and Experience*, 33(9), pp. 1–16. doi: 10.1002/cpe.6156.

Shaikh, T. A., Rasool, T. and Lone, F. R. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, 198, p. 107119.

Yang, B. *et al.* (2022). ST-FVGAN: filling series traffic missing values with generative adversarial network. *Transportation Letters*, 14(4), pp. 407–415.