Towards Temporal Dependency Identification based on Multivariate Time Series IIoT Data.

Abubakar Isah
Department of ICT Convergence System Engineering
Chonnam National University
Gwangju, South Korea
abubakarisah@jnu.ac.kr

Hyeju Shin
Department of ICT Convergence System Engineering
Chonnam National University
Gwangju, South Korea
sinhye102@gmail.com

Sani Ahmad Hassan
Department of Computer Science
Ahmadu Bello University
Zaria, Nigeria
myynazee@gmail.com

Seugmin Oh
Department of ICT Convergence System Engineering
Chonnam National University
Gwangju, South Korea
osm5252kr@gmail.com

Jinsul Kim
Department of ICT Convergence System Engineering
Chonnam National University
Gwangju, South Korea
jsworld@jnu.ac.kr

***Abstract —*In multivariate time series data, the analysis of temporal dependencies is crucial for comprehending complex relationships, forecasting trends, and streamlining diverse applications. In this study, we proposed a method for improving multivariate time series data using temporal dependence modeling. Our approach is based on utilizing cutting-edge machine learning methods, particularly Long Short-Term Memory (LSTM) networks, that are well known for their capacity to record complicated temporal correlations. We want to reveal hidden patterns and dependencies within multivariate time series data by employing LSTM networks. We use benchmark metrics Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Mean Squared Error (MSE) to assess the effectiveness of our strategy. We run comprehensive tests on various datasets to demonstrate the accuracy and robustness of our model.**

Keywords—multivariate, LSTM, temporal dependency

1. Introduction

Temporal dependencies comprise the relationships and correlations that develop over time among variables in multivariate time series data. These dependencies cover things like trends, delayed associations, sequential arrangements, seasonality, and other temporal configurations [1]. In a variety of industries, including weather forecasting, industrial operations, and countless other fields, temporal relationships and the study of multivariate time series are vital [2]. The capacity to effectively analyze and forecast [3] complex dynamics is crucial for achieving optimal decision-making and system optimization. Time-series data provides a way to examine a variety of factors, including household energy, traffic congestion, currency value variations, solar energy output, and even musical notation. The data gathered most frequently consists of multivariate time series (MTS) data, much as how a utility company may track the power use of numerous consumers. The complex and changing connections between many series [4] might be important, but observing and understanding them can be challenging. The systems that people use are getting more complicated as science and technology advance. As a result, using the LSTM networks to analyze temporal connections within multivariate time series and the idea of digital twins together create a potent method of comprehending and forecasting complex systems. These methods offer fresh opportunities for boosting effectiveness, preventative maintenance, and decision-making support across numerous industries, ultimately boosting productivity, dependability, and system performance.

2. Related work

With the advancement of deep learning, an increasing number of scholars are aiming to employ deep learning techniques for modeling the analysis of Multivariate Time Series (MTS). The RNN and its diverse adaptations stand out as the quintessential sequence-based deep learning model. However, achieving convergence can be complex due to the inherent susceptibility of standard RNN models to encountering issues like vanishing and exploding gradients [5]. The use of LSTM, a method that is still widely used in many sequential models, has helped to some extent to alleviate the problem of disappearing gradients in RNN. To forecast time series patterns, [6] coupled LSTM and the conventional genetic algorithm. The best LSTM configuration was found using a genetic algorithm, which allowed for the efficient application of time series data from the petroleum industry, [7] used LSTM in the area of supply chain analysis and forecasting. When LSTM was used to measure power load within an early warning system for power compliance, remarkable results were obtained. Additionally, [8][9] investigated several power consumption time series datasets using LSTM for evaluation. The effective forecasting of carbon pricing was achieved in [10] by the combination of the random forest technique and LSTM. In order to forecast daily precipitation time series data, [11] also built a self-encoding network with LSTM roots. Another work by [12] used LSTM to analyze historical data from oil well production and provide conclusions.

However, recurrent neural networks (RNNs), a crucial subcategory of neural networks, are used in the analysis of sequence data. However, the RNNs struggle with vanishing or bursting gradients, which makes it difficult for them to solve the problem of long-term dependencies. Long short-term memory (LSTM), a particular kind of RNN, adds gate mechanisms that successfully prevent the problems of mistake disappearance or amplification during back-propagation [13] in order to alleviate this. Unlike existing methods, StemGNN [14] adopts a novel methodology to simultaneously capture inter-series correlations and temporal dependence in the spectral domain. Investigating temporal dependencies is a crucial task in order to find temporal patterns in historical time series data, taking into account time lags and gaps in the data. This is because the industrial Internet of Things (IIoT) applications' time series data [15] operate within a dynamic system that is constantly changing. By enabling the smooth incorporation of real-time raw data from the field into digital twin systems, sensor technologies play a crucial role in Industry 4.0. However, it is important to remember that [16] sensor malfunction may be caused by either intrinsic problems or external environmental conditions.

3. Experimental result

In this section, we will comprehensively outline the experimental settings that were employed for developing both the proposed model and the reference models using authentic datasets. Furthermore, a concise summary of the reference models and the criteria for evaluating their effectiveness will be presented for the purpose of comparing their performance with the proposed model.

3.1 Dataset and Preprocessing

The BATADAL dataset contains a number of sensors and actuators that monitor and regulate the water flow in a network. It includes both typical operating situations and numerous simulated cyber-physical assaults, like sensor spoofing and actuator assaults, which might impair the system's performance and perhaps cause issues with the water supply.

Utilizing the StandardScaler to normalize the input variables is the initial strategy for addressing the sparse nature of IIoT data. Standard scaling is used in this transformation to change the data's mean and standard deviation to 0 and 1, respectively. By redistributing the data distribution around 0, this scaling technique increases the efficiency of some machine learning algorithms. To scale the input data 'inputs' uniformly, the StandardScaler is used. Managing missing values in a dataset can be difficult when training machine learning models. Utilizing the SimpleImputer approach allows for the management of this issue. This strategy calls for substituting appropriate values for any missing values. The "mean" technique was chosen in this situation, which implies that missing values are replaced with the average of the values already present in the appropriate column. 3.2 Evaluation metrics

To evaluate the efficiency of the models, various indicators were utilized, encompassing mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE). In situations where prediction accuracy is frequently expressed as a percentage, employing the MAPE metric proves advantageous for quantifying prediction precision. The MAE calculates the average discrepancy between the model's forecasts and the factual data. On the other hand, the RMSE gauges the standard deviation of the model's predictive outcomes. A lower value signifies enhanced model efficacy. The explanations for these three criteria are provided below:

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{y_i - \hat{y_i}}{y_i}\right) X\ 100 \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} y_i - \hat{y_i} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})} \tag{3}$$

All three assessment standards were applied, representing the actual value of the ith instance, signifying the predicted value of the ith instance, and denoting the mean value across n samples. Our model's loss function is the Mean Absolute Error (MAE), as detailed in **Section 3.1**. The training of our model was conducted using the Adam optimizer in combination with the LSTM model.

Algorithm: Training Process

Require Epoch, TimeSeries, Target, Length of delay, period.

for *t in range(1,len(TimeSeries)):*

Z_p = *TimeSeries*$[\max(0, t - l_d)]: t$

Z_p = *TimeSeries*$[\max(0, t - l_p * p)] :: p$

Instance = $(Z_d, Z_p, Target[t-1])$

ϑ = initialize_parameter()

for *iteration in range(NumIteration):*

random_instances=random_selection(Dataset,batch_size)

for Z_d, Z_p *Label in random_instances:*

Prediction =TD_LSTM_model_forward(Z_d, Z_p, θ)

Loss = compute_loss(Prediction, Label)

ϑ = optimize_parameters(ϑ, loss)

Trained_Model = TD_LSTM_model_with_parameters()

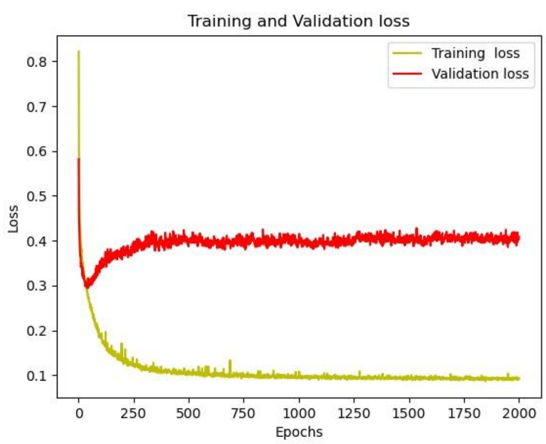

Figure 1: Training and Validation loss

Discussion

We compare our model with some other baseline models. The reference models encompassed DA-ConvLSTM [13], StemGNN [23], and LSTNet [1]. The one-dimensional CNN was employed for handling time series data, while LSTM [11] represented the conventional sequential model, and GRU [13] denoted an enhanced variant of LSTM possessing fewer parameters. These models were integrated with a dense layer featuring a solitary hidden unit. DA-ConvLSTM [13] stands as a recent entrant in the realm of MTS prediction models. This model incorporates dual attention layers and a Conv. layer to effectively capture both temporal and spatial correlations within the MTS. Notably, it has demonstrated impressive performance. For our experimentation, we adopted the training configurations as outlined by the original authors. LSTNet [1] integrates the lightweight LSTM to capture immediate and prolonged dependencies in a concerted manner. All the baseline was trained using the NASDAQ100 dataset while ours used the BATADAL dataset. We were able to improve our system by capturing the temporal dependency and evaluating it using the three evaluation metrics.

Table 1: Model comparison with other models in the literature.

Methods	BATADAL		
	MAPE	MAE	RMSE
DBN-LSTM [2]	1.1373	0.5583	0.6936
LSTNet [17]	0.2325	0.1133	0.1615
DA-ConVLSTM [6]	0.1958	0.0962	0.1416
TSA-Conv-LSTM [5]	0.1549	0.0763	0.1216
Ours	**0.0153**	**0.6952**	**1.5140**

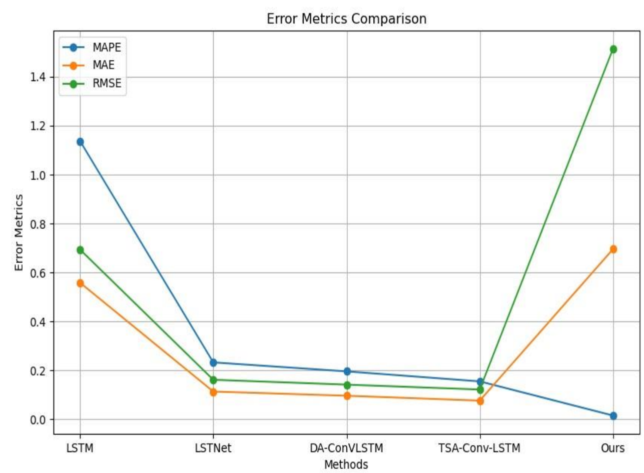

Figure 2: Error metrics comparison using MAPE, RMSE, and MAE.

We carried out extensive studies utilizing benchmark measures including Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Mean Squared Error (MSE) to confirm the effectiveness of our technique. These metrics allowed for a thorough examination of the effectiveness of our model.

4. Conclusion

This study introduces a novel approach to enhance multivariate time series data analysis by leveraging the power of temporal dependence modeling. Our methodology revolves around the application of state-of-the-art machine learning techniques, specifically Long Short-Term Memory (LSTM) networks, renowned for their ability to capture intricate temporal correlations. Through the utilization of LSTM networks, we aim to uncover concealed patterns and interdependencies embedded within the multivariate time series data. The significance of our proposed method lies in its potential to revolutionize the way we comprehend and exploit multivariate time series data. By harnessing the capabilities of LSTM networks, our approach offers a pathway to better understand complex dynamics, forecast trends, and optimize various applications.

However, there are still problems that we can enhance our system.

1. The temporal dependency remains highly reliant on LSTM units, indicating that the model's complete parallelization is not feasible.

ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Innovative Human Resource Development for Local Intellectualization support program(IITP-2023-RS-2022-00156287) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation)" and This work was partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government(MSIT) (No.2021-0-02068, Artificial Intelligence Innovation Hub).

REFERENCES

[1] G. Lai, W. C. Chang, Y. Yang, and H. Liu, "Modeling long- and short-term temporal patterns with deep neural networks," *41st Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval, SIGIR 2018*, no. July, pp. 95–104, 2018, doi: 10.1145/3209978.3210006.

[2] R. Vohra, K. Goel, and J. K. Sahoo, "Modeling temporal dependencies in data using a DBN-LSTM," *Proc. 2015 IEEE Int. Conf. Data Sci. Adv. Anal. DSAA 2015*, no. 3, pp. 1–4, 2015, doi: 10.1109/DSAA.2015.7344820.

[3] H. Widiputra, A. Mailangkay, and E. Gautama, "Multivariate CNN-LSTM Model for Multiple Parallel Financial Time-Series Prediction," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/9903518.

[4] Y. Lian, Y. Geng, and T. Tian, "Anomaly Detection Method for Multivariate Time Series Data of Oil and Gas Stations Based on Digital Twin and MTADGAN," *Appl. Sci.*, vol. 13, no. 3, 2023, doi: 10.3390/app13031891.

[5] E. Fu, Y. Zhang, F. Yang, and S. Wang, "Temporal self-attention-based Conv-LSTM network for multivariate time series prediction," *Neurocomputing*, vol. 501, pp. 162–173, 2022, doi: 10.1016/j.neucom.2022.06.014.

[6] S. Y. Shih, F. K. Sun, and H. yi Lee, "Temporal pattern attention for multivariate time series forecasting," *Mach. Learn.*, vol. 108, no. 8–9, pp. 1421–1441, 2019, doi: 10.1007/s10994-019-05815-0.

[7] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, 2019, doi: 10.1016/j.neucom.2018.09.082.

[8] K. J. Uribe-Murcia, Y. S. Shmaliy, C. K. Ahn, and S. Zhao, "Unbiased FIR Filtering for Time-Stamped Discretely Delayed and Missing Data," *IEEE Trans. Automat. Contr.*, vol. 65, no. 5, pp. 2155–2162, 2020, doi: 10.1109/TAC.2019.2937850.

[9] P. S. Kam and A. W. C. Fu, "Discovering temporal patterns for interval-based events," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1874, pp. 317–326, 2000, doi: 10.1007/3-54044466-1_32.

[10] D. Ponnoprat, "Short-term daily precipitation forecasting with seasonallyintegrated autoencoder," *Appl. Soft Comput.*, vol. 102, p. 107083, 2021, doi: 10.1016/j.asoc.2021.107083.

[11] M. Lavallee *et al.*, "Evaluating the performance of temporal pattern discovery: a new application using statins and rhabdomyolysis in OMOP databases," *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, pp. 1–7, 2022, doi: 10.1186/s12911022-01765-1.

[12] M. Randles, D. Lamb, E. Odat, and A. Taleb-Bendiab, "Distributed redundancy and robustness in complex systems," *J. Comput. Syst. Sci.*, vol. 77, no. 2, pp. 293–304, 2011, doi: 10.1016/j.jcss.2010.01.008.

[13] Y. Qin, D. Song, H. Cheng, W. Cheng, G. Jiang, and G. W. Cottrell, "A dualstage attention-based recurrent neural network for time series prediction," *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 0, pp. 2627–2633, 2017, doi: 10.24963/ijcai.2017/366.

[14] C. Özden and M. Bulut, "Spectral temporal graph neural network for multivariate agricultural price forecasting," *Cienc. Rural*, vol. 54, no. 1, pp. 1– 13, 2024, doi: 10.1590/0103-8478cr20220677.

[15] A. Isah *et al.*, "A Data-Driven Digital Twin Network Architecture in the Industrial Internet of Things (IIoT) Applications," vol. 6, no. 2.

[16] A. Aksoy, Y. E. Ertürk, S. Erdoğan, E. Eyduran, and M. M. Tariq, "Estimation of honey production in beekeeping enterprises from eastern part of Turkey through some data mining algorithms," *Pak. J. Zool.*, vol. 50, no. 6, pp. 2199– 2207, 2018, doi: 10.17582/journal.pjz/2018.50.6.2199.2207.

[17] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons and Fractals*, vol. 135, p. 109864, 2020, doi: 10.1016/j.chaos.2020.109864.