The Indonesian Journal of Computer Science
www.ijcs.net
Volume 13, Issue 4, August 2024
https://doi.org/10.33022/ijcs.v13i4.4205

Performance of Cloud Computing Resources Allocations SGLA Model Compared to ARIMA

Sello Prince Sekwatlakwatla[1], Vusumuzi Malele[2]
sek.prince@gmail.com[1],vusi.malele@nwu.ac.za[2]
[1]Unit for Data Science and Computing School of Computer Science and Information Systems North-West
[2]University Vanderbijlpark, South Africa

Article Information

Received : 26 Jun 2024
Reviewed : 19 Jul 2024
Accepted : 1 Aug 2024

Keywords

Resource allocation, Traffic flow, Ensemble method

Abstract

Solutions for cloud computing are growing in popularity as a means for businesses to streamline operations, save expenses and increase productivity. The benefit of cloud services is that they let customers access on-demand apps and services from a shared pool of programmable computer resources and store data offsite. Cloud computing resource allocation requires sophisticated tools and methodologies for optimal utilization. These problems include load balancing, efficient resource management and compliance with legal and regulatory requirements. The majority of businesses are switching to cloud services and advising their customers to use internet services. However, effective resource allocation is critical for improving performance and lowering costs in this area. Due to unpredictable network traffic in cloud computing, resource allocation is challenging, which causes customers to complain about application timeouts, delayed system times and higher bandwidth use during peak hours. This entails allocating resources to various users and programs, such as memory, processing power, storage, and network bandwidth. In this regard, this study compares the performance ensemble method, which is stepwise Gaussian Linear Autoregressive (SGLA), and the individual method which is autoregressive integrated moving average (ARIMA). The Matlab tool is used for simulation and evaluation of the results. The results show SGLA prediction accuracy increased to an average of 98.9%, and ARIMA prediction showed an accuracy of 75.5%. In this regard, the ensemble method performed better than individual methods using the same datasets. The study recommends the ensemble method for the prediction and allocation of resources in cloud computing.

A. Introduction

Cloud computing facilitates speedier innovation, flexible resource allocation and economies of scale by offering services such as servers, storage, databases, networking, software, analytics and intelligence via the internet[1-2]. Cloud computing adapts to business growth, reducing waste and security threats. It enables faster data storage, extraction and transformation, enhancing efficiency in online living and automated manufacturing, thanks to encryption, access limits, and monitoring.For businesses to succeed in their digital transformation efforts and be able to grow and adjust to changing market demands, cloud computing is essential. It provides consumers with flexible, on-demand access to technology, hence promoting digital service offerings and operations[3]. Without requiring local hardware or infrastructure, cloud computing provides on-demand computing services, such as power, storage, and applications, over the internet. Through servers in data centers that are maintained, secured and updated by cloud service providers, users may access resources remotely[4].

The majority of organizations adopt cloud integration because it meets technological needs, promotes expansion and agility and makes collaboration between IT and other business divisions more seamless however, most service providers are receiving complaints from their customers regarding slow system speed, application timeouts and the inability to access certain services[4-5]. Scalability issues may prevent genetic algorithms from being used in real-world situations[6].

Utilizing cloud parallelization might be an inexpensive means of obtaining timely solutions that meet the requirements of fault tolerance, resource discovery, scalability and cost-effectiveness. However, there's a fair probability that the communication cost associated with parallel computing will be substantial, and it's difficult to distribute genetic algorithms on demand[7]

An adaptive ensemble method is proposed for identifying server abnormalities in cloud computing; the method lowers complexity and increases accuracy in traffic flow; however the method's limitation is that a high volume of resources is available even if the cloud traffic is low[8]. Clouds allow companies to rent resources for storage and computational purposes, thereby significantly reducing their infrastructure costs. To increase the accuracy of predicting the volatility of crude oil futures prices, the event-driven gated recurrent unit technique is suggested [9]. Nevertheless, it may have adaptivity, generalization, overfitting and narrow decision space problems.

By applying hybrid particle swarm techniques, a support vector machine was developed to increase the accuracy of least squares optimization of high-speed trains[10].For turbofan engine predictive maintenance strategy prediction, a deep learning ensemble approach is suggested. Experimental results demonstrate that the technique is efficient and better than current approaches and that the method can save costs and increase dependability while taking the engine's mission cycle and maintenance operations into account. However additional possibilities were not taken into consideration and the study solely concentrated on engine mission cycles. In order to customize maintenance tactics for various mission cycles, future research should take decision-maker preferences, dynamic strategies and RUL prediction techniques into account [11]. In this study multiple linear regression

(MLR) and artificial neural network (ANN) is proposed and models included with outstanding prediction performance seen in the findings. Nevertheless, it has been challenging to accurately capture the motion characteristics of the vessel in response to the sea's weather using in-service data gathered at one-minute intervals [12]. Research on these algorithms for financial risk prediction has shown that the convolutional neural network method improves accuracy and reliability by converting financial data into two-dimensional image data; risk analysis is not taken into account[13].

In an effort to identify the best-enhancing model, a number of research have been done in the field of ensemble techniques [14]. The majority of the findings indicated that the model would perform well enough to be used in a real-time system. As an instance, in [15] Sector errors are addressed by a suggested ensemble pruning-based prediction technique for cloud storage systems, which leads to improved accuracy, shorter training times and less memory use. a three-stage stochastic programming methodology was used to reduce waiting times, emergency transportation expenses and unfulfilled demand in relief organizations during peak hours. However, resource allocation declined significantly, especially with high traffic volume[16]. Consequently, a number of novel ideas and techniques for allocating resources have been proposed.Stepwise Gaussian Linear Autoregressive (SGLA) is an ensemble model that [17] proposed to be used in their study. The purpose of this study is to validate SGLA's performance in this regard by contrasting it with the autoregressive integrated moving average(ARIMA). The research question "Can the SGLA model perform better than ARIMA model?" The remainder of the paper is organized into sections that include research method, result and discussion, conclusion and future work.

B. Research Method

1. ARIMA

Moving average (MA) and autoregressive (AR) components are used in the Autoregressive integrated moving average(Arima) approach for signal strength forecasting. ARIMA is a statistical model that uses time series data to improve knowledge of a data collection or anticipate future trends [18]. In this sense, Arima may be used with or without regressors. The study uses Arima to forecast resource allocation. This is how the typical Arima formula is explained:

$$y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$

where the data to be subjected to the ARMA model is denoted by {yt}. Thus, the series has already undergone power transformation and difference, in that sequence.AR coefficients are represented by parameters φ1, φ2 and so on.

The formula for a moving average (MA) model of order q, or MA (q) model:

$$y_t = C + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2)$$

Where {yt} is defined as before and φ1, φ2 and so on are MA coefficients.

An ARMA (p,q) model equation:

$$y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3)$$

$\{y_t\}$, φ_1, φ_2..., θ_1, θ_2... are as defined before.

Formula for a SARMA (p,q) (P,Q) model (seasonal):

$$y_t = C + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{P} \Phi_i y_{t-is} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \sum_{i=1}^{Q} \Theta_i \varepsilon_{t-is} \quad (4)$$

$\{yt\}$, $\{\varphi\}$, and $\{\theta\}$ are as defined before, and $\{\Phi\}$ and $\{\Theta\}$ are the seasonal counterparts.

The constant component in an ARIMA equation introduces and extends a deterministic trend, with linear trends for one difference and quadratic trends for two differences.The model calculates the value of a constant term when it is included.

$$C = \mu \times (1 - \sum \phi_i)(1 - \sum \Phi_i) \quad (5)$$

φi are nonseasonal AR coefficients, Φi are seasonal AR coefficients, and μ is the mean of the series.

2. SGLA

As a technique for resource allocation using data analytics, the proposed method, called SGLA, combines the use of linear regression, support vector machines, Gaussian process regression, and the autoregressive integrated moving average [17]. Stepwise regression, forward selection, and backward elimination are the three forms of the hierarchical stepwise regression technique that are known [19]. Statistical algorithms are used to identify predictors as the model is built. These algorithms start with all possible predictors and gradually add or eliminate them.

$$Pj.std = bj\left(\frac{fx_j}{fy}\right) \quad (6)$$

Where fy ,fx_j is the normal deviations for the reliant variable and the consistent Pj independent variable

- **Support vector machines**

Support Vector Machines (SVMs) efficiently use a subset of training points in decision functions for classification, regression, and outlier detection, with predictions explained by Equations 2 and 3[19].

$$W(c_i) = \text{sign}\left(\sum_{j=1}^{s} \propto_j p_j k(c_{j,} x_i) + b\right) \quad (7)$$

$$W(v, v^1) = \exp\left(\frac{\|v - v^1\|^2}{2y^2}\right) \quad (8)$$

Where ci is the (vector of values) to predict. The $c_{j,}$se are called support vectors which are a subset of the training data. The p_j re is the class (-1 or +1) of each data p_j. The $c_{j,}$re are constants, one for eachp_j. The b is a single numeric constant. Letter s is the number of support vectors. The W is a kernel function returns a number, 1.0 meaning identical and 0.0 meaning as different as possible, based on the similarity between two vectors.

- **Gaussian Process Regression**

A finite number of random variables with a Gaussian distribution make up Gaussian processes. Which are input, Gaussian field, and observation. The model's conventional formula is as follows[19].

$$J(x) = \sum_{i=1}^{n} \beta_i f_i(x) + F(x) \quad (9)$$

Where Ji is the reaction, we are interested in, $F(x)$ is a Gaussian process, $f\,i$ is known functions, and β is unknown. Suppose there are n sample points x1, x2…, xm,with corresponding sample results y1,y2,…,ym, β can then be estimated using the equation. Therefore, we can use it.

$$J(X) = \sum_{i=1}^{n} \beta_i f_i(X) \quad (10)$$

Then forecast response at sites x.

As various machine learning contests have shown, ensemble approaches in machine learning integrate multiple models for better outcomes, frequently producing more accurate solutions than a single model. In this regard, the below figure shows the SGLA method (see figure 1).

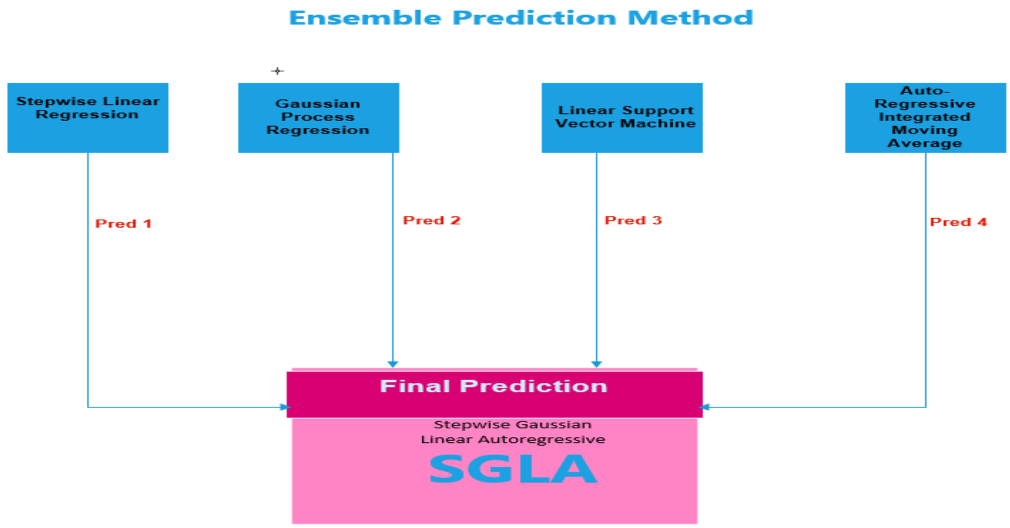

Figure 1. Research design Method [17]

3. Indicators for model assessment

This work proposes criteria for evaluating projected accuracy, such as root mean square deviation (Rmsd) and mean absolute percentage error (MAPE), to test the validity of the SGLA model [20].

$$Mape = \frac{100}{n} \sum_{t=1}^{n} \frac{O_t - W_t}{O_t} \quad (11)$$

Where O_t is the actual value and W_t is the predicted value.

$$\text{Rmsd} = \sqrt{\frac{\sum_{i=1}^{R} (H_t - \widehat{H_i})^2}{A}} \quad (12)$$

Where t = variable 1, A Number of data points with no missing values, H_t is the real observations time series and $\widehat{H_1}$ the time series is estimated [21].

C. Result and Discussion

1. SGLA

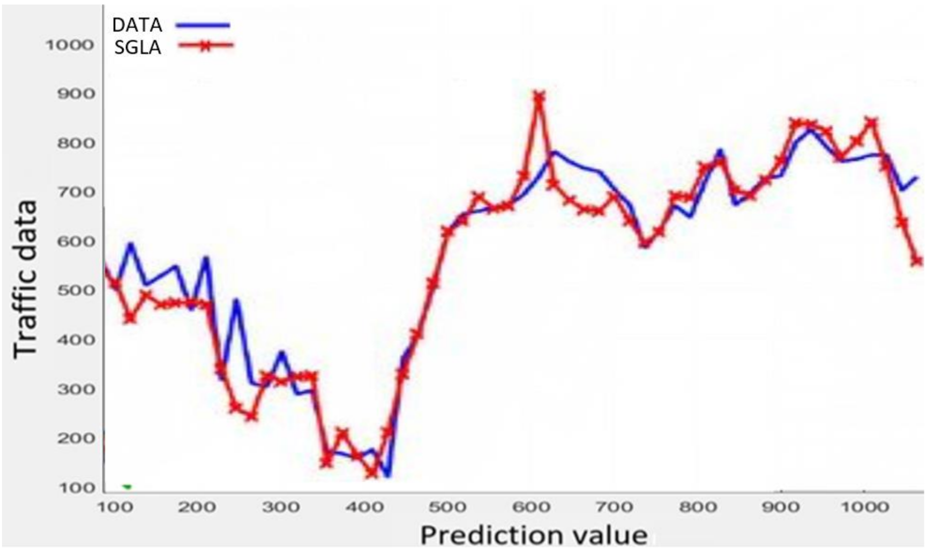

Figure 1. Results of SGLA

In Figure 2, SGLA prediction is shown: the accuracy reduces with decreasing data; on the other hand, the accuracy grows with rising traffic, the technique performs better, and the allocation accuracy increases with an average of 98.9% (see table 1).

Table 1. SGLA Results

	Results Analysis		
Traffic Data (%)	**Resource Allocation (%)**	**Mape (%)**	**Rmse (%)**
59	60	1.73	1.00
49	48	2.13	1.00
56	55	1.83	1.00
37	38	2.88	1.00
12	13	9.10	1.00
33	34	0.00	0.00
62	62	0.00	0.00
99	75	18.69	17.00
65	65	0.00	0.00
72	73	1.40	1.00
63	64	1.40	1.00
73	75	2.89	2.00
70	70	0.00	0.00
66	67	1.40	1.00
56	66	18.18	10.00

Table 2 provides a summary of how the resources are distributed. The distribution of the data was superb, especially as the amount of traffic increased. The analysis revealed that its accuracy was 98.9% overall. The maximum error for

RMSE was 17%, the lowest Mape error was 0%, and the largest Mape error was 18.69%. This strategy produced the most accurate forecast with 0% error.

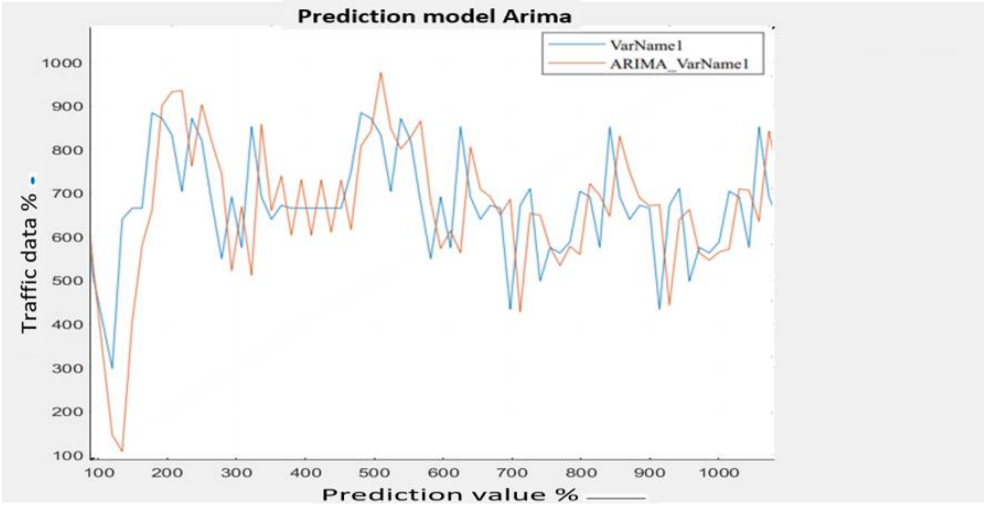

Figure 2. Results of ARIMA

The Autoregressive integrated moving average is displayed in Figure 3. The methodology forecast accuracy is 75.5%, and as traffic rises, the model's prediction becomes more inaccurate.

Table 2. ARIMA Results

	Results Analysis		
Traffic Data (%)	**Resource Allocation (%)**	**Mape (%)**	**Rmse (%)**
62	53	15.01	9.00
34	37	9.09	1.00
12	31	63.34	19.01
90	93	3.38	3.00
86	96	11.76	10.00
73	77	5.56	4.00
66	68	3.08	2.00
92	75	18.69	3.00
84	81	3.62	3.00
86	84	2.36	3.00
44	50	8.89	4.00
53	43	19.24	10.00
58	60	3.52	2.00
87	82	5.75	5.00
70	61	12.86	9.00

Table 2 provides a summary of how the resources are distributed. As traffic volume grows, the forecast accuracy decreases. The analysis revealed that, on the whole, its accuracy was 75.5%. The maximum error for RMSE was 19.1%, the lowest Mape error was 2.36%, and the largest Mape error was 63.34%. This approach produced the best accurate forecast with a 1% error.

D. Conclusion

Two prediction techniques were examined in this study: ARIMA prediction shown an accuracy of 75.5%, whereas SGLA prediction accuracy rose with an average of 98.9%. Using the same datasets, the ensemble approach outperformed individual techniques in this aspect. The ensemble approach is suggested by the study for resource allocation and prediction in cloud computing. The largest challenge is getting resource allocation findings that are 100% correct when compared to actual data. Therefore, additional investigation is still needed into the accurate traffic forecast in cloud computing systems. Additionally, study is needed to ascertain the system specs and server count necessary to connect to and host this solution.

E. Acknowledgment

The article is a component of the PhD program in Computer and Information Sciences with Information Technology at North-West University.

F. References

[1] J. Ren, D. Fu, C. Shi, Z. Huang, W. Zhu and Y. Liu, Research on Cloud Computing Technology Graph Analysis, 2023 2nd International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT), Xiamen, China, pp. 84-91,2023.

[2] N.C.Galego, D.S.Martinho, N.M.Duarte,Cloud computing for big data analytics How cloud computing can handle procesing large amounts of data and improve real-time data analytics,Journal of Procedia Computer Science,Vol.237.Pp 297-304,2024,[Online].Available:https://doi.org/10.1016/j.procs.2024.05.108

[3] D. H. Nam, "A Comparative Study of Mobile Cloud Computing, Mobile Edge Computing, and Mobile Edge Cloud Computing," 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), Las Vegas, NV, USA, pp. 1219-1224, 2023.

[4] H.Ma,ADing,"Method for evaluation on energy consumption of cloud computing data center based on deep reinforcement learning",Journal of Electric Power SystemsResearch,Vol,208,2022,[Online].Available:https://doi.org/10.1016/j.epsr.2022.107899

[5] F. Wang, J. -t. Zhou, H. Wang and X. Guo, "A Blockchain-Based Multi-Cloud Storage Data Consistency Verification Scheme," 2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), Melbourne, Australia, pp. 371-377,2022.

[6] C. Xia, R. Wang, Z. Deng and Y. Zheng, "Research on cloud computing resource scheduling based on improved ant colony optimization algorithm," 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qingdao, China, pp. 295-298,2022

[7] L.Cai, C.Xing, Y.Deng,"Research on digital urban architecture design based on cloud computing data center",Journal of Environmental Technology &

Innovation,Vol.22,2021,[Online].Available:https://doi.org/10.1016/j.eti.2021.101543
[8] L. Ruan et al., "Cloud Workload Turning Points Prediction via Cloud Feature-Enhanced Deep Learning," in IEEE Transactions on Cloud Computing, vol. 11, no. 2, pp. 1719-1732,2022
[9] K. Ramesh, A. Lakshna and P. N. Renjith, "Smart Traffic Congestion model in IoT-A Review," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, pp. 651-658,2020.
[10] J.S.Pimentel, R.Ospina, A.Ara,"A novel fusion Support Vector Machine integrating weak and sphere models for classification challenges with massive data",Journal of Decision Analytics Journal,Vol.11,2024,[Online].Available:https://doi.org/10.1016/j.dajour.2024.100457
[11] Y.Chen, C.Xiao, S.Yang, Y.Yang, W.Wang,"Research on long term power load grey combination forecasting based on fuzzy support vector machine",Journal of Computers and Electrical Engineering,Vol.116,2024,,[Online].Available:https://doi.org/10.1016/j.compeleceng.2024.109205
[12] K. Khan, S. Husain, G. Nauryzbayev and M. Hashmi, "Development and Evaluation of ANN, ACOR-ANN, ALO-ANN Based Small-Signal Behavioral Models for GaN-on-Si HEMT," 2023 30th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Istanbul, Turkiye, pp. 1-4, 2023
[13] K. S. Aswin, M. Purushothaman, P. Sritharani and A. T. S, "ANN and Deep Learning Classifiers for BCI applications," 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), Kannur, India, 2022, pp. 1603-1607,2022
[14] N. Ramteke and P. Maidamwar, "Cardiac Patient Data Classification Using Ensemble Machine Learning Technique," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, pp. 1-6,2023.
[15] M. Rahman, M. I. A. Palash, M. S. Sumi, M. Rahman and M. J. Al Nahian, "Using Ensemble Approaches and Different Sampling Techniques to Handle Class Imbalance Challenges in Coronary Heart Disease Prediction," 2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), Dhaka, Bangladesh, 2023, pp. 31-35,2023
[16] M. Shah, K. Gandhi, K. A. Patel, H. Kantawala, R. Patel and A. Kothari, "Theoretical Evaluation of Ensemble Machine Learning Techniques," 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, pp. 829-837, 2023
[17] S.Sekwatlakwatla & V. Malele, "Model for Enhancing Cloud Computing Resource Allocation Management Using Data Analytics"Journal of Information Systems and Informatics, vol.6, 514-526. [Online].Available: https://journal-isi.org/index.php/isi/article/view/679
[18] A. Gupta and A. Kumar, "Mid Term Daily Load Forecasting using ARIMA, Wavelet-ARIMA and Machine Learning," 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and

Commercial Power Systems Europe (EEEIC / I&CPS Europe), Madrid, Spain, pp. 1-5,2020.

[19] S. Chen, R. Lin and W. Zeng, "Short-Term Load Forecasting Method Based on ARIMA and LSTM," 2022 IEEE 22nd International Conference on Communication Technology (ICCT), Nanjing, China, pp. 1913-1917,2022.

[20] X. Chang, "Soft Measurement Modeling of Turbidity in Flocculation Process of Drinking Water Treatment Using Gaussian Process Regression," 2021 40th Chinese Control Conference (CCC), Shanghai, China, pp. 6196-6200,2021.

[21] Al-Khowarizmi, S. Efendi, M. K. Nasution and M. Herman, "The Role of Detection Rate in MAPE to Improve Measurement Accuracy for Predicting FinTech Data in Various Regressions," 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia, pp. 874-879,2023.