Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 1 Iss. 2 January 2024, pp: 85-94
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Performance Comparison of ARIMA, LSTM and SVM Models for Electric Energy Consumption Analysis

Cintia Putri Trisya[1], Nilam Wahdiaz Azani[2*], Laras Mayangda Sari[3] , Hani Handayani[4], Muhammad Rizki Miftha Alhamid[5]

[1,2,3,4]Department of Information System, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[4]Department of Computer Science, San Jose State University, USA

E-Mail: [1]12050320325@students.uin-suska.ac.id, [2]12050320451@students.uin-suska.ac.id, [3]12050321657 @students.uin-suska.ac.id, [4]12050321652@students.uin-suska.ac.id, [5]muhammadrizkimiftha.alhamid@sjsu.edu

Received Aug 6th 2023; Revised Oct 3rd 2023; Accepted Dec 11th 2023
Corresponding Author: Nilam Wahdiaz Azani

Abstract

The changing needs of electrical energy result in the electrical power needed for everyday life being unstable, so planning and predicting how much electrical load is needed so that the electricity generated is always of good quality. So it is necessary to predict the consumption of electrical energy by using forecasting on the machine learning method. Support Vector Machine (SVM), Autoregressive Integrated Motion Average (ARIMA), and Long Short-Term Memory (LSTM) are models that are often used to overcome patterns in predictions. To find out the best models how to predict electricity consumption in the future and how the SVM, LSTM, and ARIMA algorithms perform in predicting electricity consumption. This research will look for the RMSE value and prediction time, then compare it with the best average value. The results of the study show that the ARIMA model is able to predict electricity usage for the next 1 year period, in the evaluation using the RMSE metric, where SVM shows a much lower value than ARIMA and LSTM. In this case, SVM achieved RMSE of 0.020, while ARIMA and LSTM achieved RMSE of 7.659 and 11.4183, respectively. Even though SVM has a lower RMSE, it is still unable to predict electricity usage for the next 1 year with sufficient accuracy.

Keyword: ARIMA, Electrical Energy, Machine Learning, SVM, LSTM, RMSE

1. INTRODUCTION

Uncertain changes in the demand for electrical energy result in fluctuations in the demand for electric power, so it is necessary to plan, regulate, and estimate the electric load to ensure that the production of electric power remains consistent with good quality [1] . The use of electrical energy has a very important and urgent role throughout the world. In this day and age, no human being can escape the use of electricity. Many activities use electrical energy, including activities in households, offices, electronics, and many more [2]. Therefore electricity consumption has increased every year. Therefore, for a country or region, the projection of the use of electrical energy is urgent and critical [3]. Among the various sectors mentioned, the industrial sector has a relatively high level of energy consumption compared to other sectors. Therefore, efforts are needed to regulate energy consumption, especially in the industrial sector.

Prediction of electric load has a significant contribution in carrying out the operation of the electric power system, including generation planning, power flow analysis, unit commitment, hydro-thermal system, and economical operation of the power system [4]. The need for electrical energy on a daily or even hourly scale tends to fluctuate, so the State Electricity Company (PLN) or a provider of electrical energy needs to be able to predict the need for electrical loads every day. There are various methods that can be used to predict electricity needs, and it is very important to choose a reliable method with a high degree of accuracy to optimize the cost of producing electrical energy.

Based on the explanation above, it is necessary to predict the consumption of electrical energy using forecasting on the machine learning method. Machine Learning is a branch of Artificial Intelligence that focuses on developing systems capable of automatically learning. In this context, learning refers to the process of recognizing intricate patterns within vast amounts of data. Truly learning machines are algorithms that review data and are able to predict future behavior [5] .

In this study also used statistical methods where this method is widely used. By adopting a mathematical approach based on historical data, this method does not require time-consuming computations so that it remains a good choice with a high degree of accuracy [6] .

Several previous studies have used the Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) modeling methods to produce good accuracy regarding electricity consumption. In a study conducted by Adhib Arfan and Lussiana ETP (2020), the LSTM algorithm with SVR was applied. The research results show that LSTM is able to overcome long-term time dependence and accurately predict stock prices. [7]. According to Rifando and friends (2022) in his research, it was stated that the LSTM + Gated Recurrent Unit (GRU) algorithm showed the highest accuracy by R-squared 94% [8] .

Furthermore, research conducted by Titik Rahmawati and friends stated that the ARIMA model (0,1,0) with the Mean Absolute Percentage Error (MAPE) value of this model was 0.33%, which resulted in an accuracy of 99.67%. [9] . In their research, Mohamad Ilyas Abas and Irawan Ibrahim (2019) applied the SVM Particle Swarm Optimization (PSO) Optimization approach to predict electricity consumption . The final results of the study concluded that SVM- PSO has the ability to predict time-series data with minimal error [10]. A study conducted by Ilham Amarulloh (2021) presented findings that the ARIMA model (2,1,2) produces a Mean Square (MS) of 82.0017. Comparative analysis between the data before and after forecasting shows that the value of the data after the forecasting process is higher than the value before forecasting. [11] .

In this study, researchers want to know how to predict electricity consumption in the future and how the performance of the SVM, LSTM, and ARIMA algorithms predicts electricity consumption. This research is expected to provide benefits which include, among other things, (1) providing information regarding projections or predictions of future use of electrical energy; (2) becoming an evaluation for the public/ consumers so they can pay attention to future electricity usage, and (3) obtaining the algorithm with the best performance in predicting future electricity consumption.

2. MATERIAL AND METHOD

In this study, the methods used can be found in Figure 1. In the first step, a literature review was carried out where the researcher collected information from various sources such as books or journals as a reference in preparing the research. then collect electrical energy data, then do pre-processing and cleaning on the data that has been collected. Then, predictions are made using the LSTM, ARIMA and SVM algorithms. After that the result is a prediction of electricity usage in 2024. In the final stage, a report is prepared in which the results of the discussion of the research that has been carried out are recorded and explained in detail.

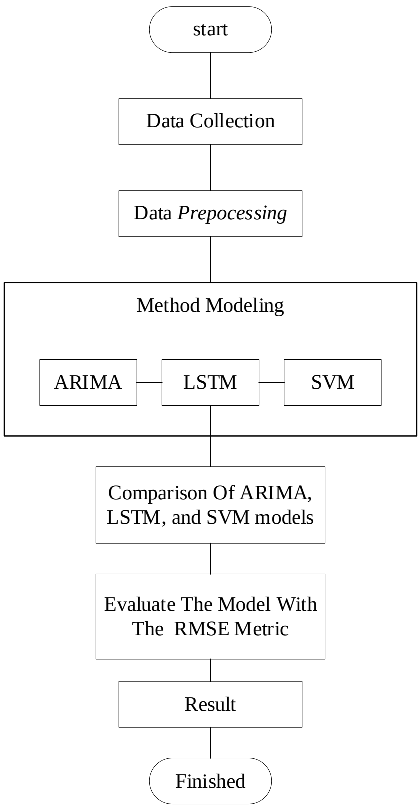

Figure 1. Research Stages

2.1 Datasets

In this research, the dataset used comes from public sources which can be accessed via the Kaggle website. This dataset is time series data covering the time period from 1985 to 2018. As the next step, we randomly divided the data into two parts, namely training data and test data. This step was carried out with the aim of showing how much electricity consumption can be predicted by the model developed in this research.

2.2 Autoregressive Integrated Moving Average (ARIMA)

ARIMA is the Autoregressive Integrated Motion Average Model. It contains three parts, AR (automatic regression), I (integration), MA (moving average)[12] [13] . he ARIMA model offers several benefits such as operating in an online learning environment, having sample sizes that do not incur storage costs, and enabling scalable and efficient parameter estimation[14] . It is denoted by ARIMA(p, d, q), where p, d, and q are the order of the autoregressive, differencing, and moving average parts of the model, respectively [15].

$$\Delta\, yt = \varphi\ 1\Delta\, yt - 1 + \varphi\ 2\Delta\, yt - 2 + \varphi p\ \Delta\, yt - p + et - \theta\ 1\ et - 1 - \theta\ 2\ et - 2 - \theta pet - p \tag{1}$$

Description:

t	= Coefficient of time
Yt	= Series value at time t
$Yt - 1$, $Yt - 2$	= Series past value
$et - 1, - 2$	= Past value of residuals
$\theta\ 1$, θp , $\varphi\ 1$, φp	= Model Coefficients

2.3 Long Short Term Memories (LSTM)

LSTM is the result of the development of the RNN method. RNN (Recurrent Neural Network) is one of the various variations of Neural Networks in the field of deep learning. This method is specifically used to predict long-term dependence [16] [17] .

1. Input Gates:

$$i(t) = \sigma(Wi \cdot [h(t\text{-}1), x(t)] + bi) \tag{2}$$

$$C\_hat(t) = \tanh(Wc \cdot [h(t\text{-}1), x(t)] + bc) \tag{3}$$

2. Forget Gate:

$$f(t) = \sigma(Wf \cdot [h(t\text{-}1), x(t)] + bf) \tag{4}$$

3. Cell State:

$$C(t) = f(t) * C(t\text{-}1) + i(t) * C\_hat(t) \tag{5}$$

4. Output Gates:

$$o(t) = \sigma(Wo \cdot [h(t\text{-}1), x(t)] + bo) \tag{6}$$

5. Hidden States:

$$h(t) = o(t) * \tanh(C(t)) \tag{7}$$

Notation:

T	= Current time step
h(t)	= Hidden state at time t
x(t)	= Input at time t
i(t)	= Input gate at time t
C_hat(t)	= Candidate cell state at time t
f(t)	= Forget gate at time t
C(t)	= Cell state at time t
o(t)	= Output gate at time t
Wi, Wc, Wf, Wo	= Weight matrices for the respective input, candidate cell state, forget gate, and output gate

bi, bc, bf, bo	= Bias vectors for the respective input, candidate cell state, forget gate, and output gate
σ	= Sigmoid function
tanh	= Hyperbolic tangent function

2.3 Support Vector Machine (SVM)

SVM is a method related to Data Mining because it is a classification algorithm based on the principle of linear classification. SVM is a learning system that uses linear function theory on features that have been trained by utilizing optimal theory [18] . The basic principle of SVM is an innovation that allows linear classification to be used to process non-linear problems. The Support Vector Machine (SVM) is assisted by the Kernel Principal Component Analysis (KPCA) with a Genetic Algorithm (GA) [19] . SVM can be used to reduce feature vector dimensions, and GA is used to optimize different SVM parameters. [20] . SVM can be used to solve regression-based problems [21] .

SVM accuracy results are affected by the parameters and kernel functions used. Table 1 presents the formulas for the various types of kernel functions to choose from.

Table 1. Kernel formula in SVM

Kernel Name	Kernel Function
Linear (Dot)	G (x 1, x 2) = x' 1, x 2
Radial basis function (RBF)	$G(x_1, x_2)$ = exp
Polynomial	$G(x_1, x_2) = yx_1 x_2 + c$

SVM operates on the principle of classifying data into two groups by finding the optimal hyperplane. The following is the calculation formula for SVM:

1. Data Poínt:

$$x_i = \{ x_1, x_2, ..., x_n \} \in R^n \quad (8)$$

2. Class Labels:

$$y_i \in \{-1, +1\} \quad (9)$$

3. Data and Class Paír:

$$\{ (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) \} \quad (10)$$

4. Maxímíze the Objectíve Functíon:

$$L_d = \sum_{i=1}^{N} \alpha_i - \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (11)$$

$$\text{Subject to: } 0 \le \alpha_i \le C \text{ and } \sum_{i=1}^{N} a_i y_i = 0$$

5. Calculate the values of w and b:

$$w = \sum_{i=1}^{N} a_i y_i x_i$$

$$b = -1/2\ w \cdot x^+ + w \cdot x^- \quad (12)$$

6. Classífícatíon Decísíon Functíon: sígn(f(x))

$$f(x) = w \cdot x + b \text{ or } f(x) = \sum_{i=1}^{m} a_i y_i K(x, x_i) + b \quad (13)$$

Explanation:

N	= Number of data points
n	= Dimension of data or number of features
Ld	= Dual Lagrange Multiplier
ai	= Weight value for each data point
C	= Constant value
m	= Number of support vectors/data points that have αi > 0
K(x, xi)	= Kernel function

2.4 Conventional RMSE Calculations

N represents the number of points measured. I_i^{Meas} dan I_i^{Calc} the calculations represent the measured and estimated solar cell currents at point *i* , respectively. Pseudo substitution equation. (1) into Eq. (2); in $U_i^{Meas} = I_i^{Meas}$, one can formulate the RMSE expression as equation 14.

$$\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N} \quad (I_i^{Meas} - I_i^{Calc})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} \quad \left(I_i^{Meas} - \left(I_{pv} - I_0\left(e^{\frac{U_i^{Meas}+I_i^{Meas}\times R_S}{n\times V_{th}}} - 1\right) - \left(\frac{U_i^{Meas}+I_i^{Meas}\times R_S}{R_P}\right)\right)\right)^2}}$$

$$\left(I_{pv} - I_0\left(e^{\frac{U_i^{Meas}+I_i^{Meas}\times R_S}{n\times V_{th}}} - 1\right) - \left(\frac{U_i^{Meas}+I_i^{Meas}\times R_S}{R_P}\right)\right) \tag{14}$$

The final segment of this equation does not indicate the computed value of the current output from the solar cell, and as a result, Eq. (equation number) does not represent the actual output current. is not a true expression of the RMSE measure. It should be noted that many works in the literature have reported using this relationship to estimate the RMSE.

3. RESULTS AND DISCUSSION

The dataset used in this study is the dataset of electricity usage. This dataset is data from 1985 to 2018 taken from the Kaggle website . At this stage the data used will also be checked using Augmented Dicky.

3.1 ARIMA

Statistically, the stationarity test is performed by examining the presence of significant trends and seasonal patterns in the data. The stationarity test is required in the analysis of time series predictions using statistical models because these models can only be applied to stationary data.

Hypothesis testing:

H0 : Data is not stationary
Ha : Stationary data
Reject H0 if p-value < alpha (0.05) is obtained).

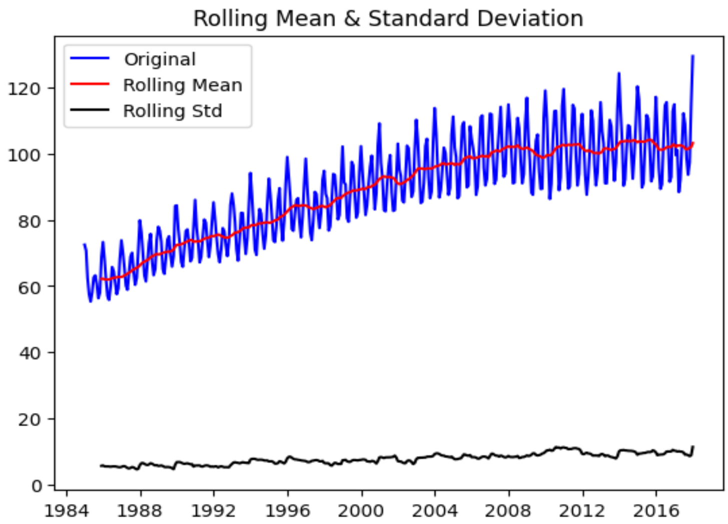

Figure 2. Dickey-Fuller Test Results

Results of Dickey-Fuller Test:

Test Statistics	: -2.256990
p-value	: 0.186215
#LagsUsed	: 15.000000
Number of Observations Used	: 381.000000

Critical Value (1%)	: -3.447631
Critical Value (5%)	: -2.869156
Critical Value (10%)	: -2.570827

From the results above, it is obtained that a p-value > 0.05 is obtained, then H0 is not rejected. So, the data has a state that is not stationary. Because the data is not stationary, it is necessary to carry out the data stationarization process, namely the differencing process to obtain stationary data.

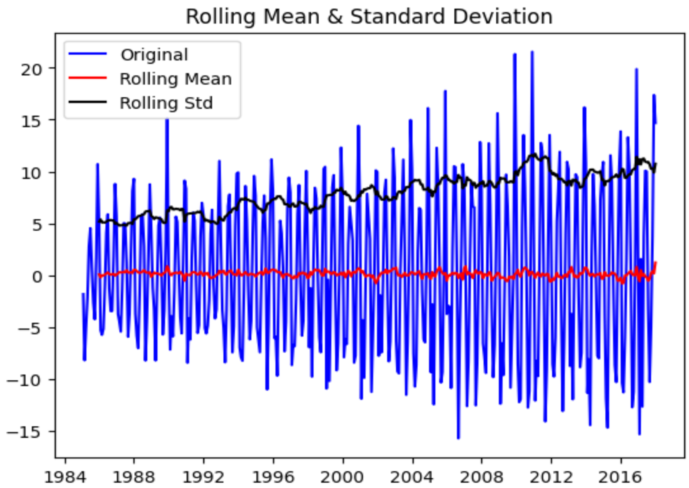

Figure 3. Rolling mean & standart deviation

Results of Dickey-Fuller Test:

Test Statistics	: -7.104891e+00
p-value	: 4.077787e-10
#Lags Used	: 1.400000e+01
Number of Observations Used	: 3.810000e+02
Critical Value(1%)	: -3.447631e+00
Critical Value(5%)	: -2.869156e+00
Critical Value(10%)	: -2.570827e+00

After the differencing process was carried out and the data stationarity was tested (after differencing), the data obtained after 1x differencing had formed stationary data. Furthermore, predictions can be made of electricity usage for the next 1 year until 2019.

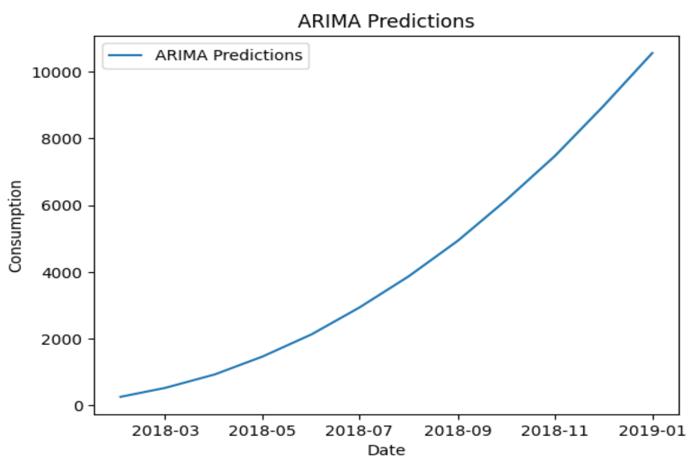

Figure 4. Arima Prediction Results

Obtained the prediction results of electricity consumption for the next 1 year in the date period, can view the table 2.

Table 2. Prediction of result 1 year

Period Year	Electricity Usage/ Kwh
2018-02-01	262.881901
2018-03-01	530.134155
2018-04-01	931.183379
…………	…………
…………	…………
2018-11-01	7484.890356
2018-12-01	8956.328973
2019-01-01	10561.566282

Table 2 displays the predicted results of electricity use in kilowatt-hours (Kwh) for a 1 year period. The time span starts from February 2018 to January 2019. This prediction provides an estimated value of electricity usage for each month during that period, which varies from 262.881901 Kwh in February 2018 to reaching the highest value of 10,561.566282 Kwh in January 2019. This data provides an overview regarding projected electricity consumption during that period, which can be used for analysis and planning of energy needs.

Table 3. Evaluation of the ARIMA model

ARIMA Evaluation	RMSE value
ARIMA(0, 0, 0)	18,542
ARIMA(0, 0, 1)	11.185
…………	………
………..	………
Best ARIMA(1, 0, 1)	7,659

In the ARIMA evaluation using RMSE values, there are several ARIMA models that are evaluated, including ARIMA (0, 0, 0), ARIMA (0, 0, 1), and ARIMA (1, 0, 1). Here are the results of each ARIMA:

1. ARIMA(0, 0, 0):
2. ARIMA model (0, 0, 0) shows an RMSE value of 18,542.
3. ARIMA(0, 0, 1):
4. The ARIMA model (0, 0, 1) shows an RMSE value of 11,185.
5. Best ARIMA(1, 0, 1):
6. The ARIMA model (1, 0, 1) gives the best results with an RMSE value of 7,659.

Based on this evaluation, ARIMA (1, 0, 1) is the best model that provides the most accurate prediction results based on lower RMSE values. However, it is important to note that this evaluation is based only on RMSE values and can be supplemented with other evaluation metrics to get a more complete picture of ARIMA model performance.

3.2 Support Vector Machine (SVM)

The figure 5 shows the evaluation results of the Support Vector Machine (SVM) method used in a scenario that involves dividing data into three groups, namely 70:10, 90:10, and 80:20. The image above does not show the movement of the desired graph. This shows that the SVM algorithm is not suitable for processing time series data. Evaluation is done by measuring the Root Mean Square Error (RMSE) value, which is a common metric used to measure how well the predictive model matches the actual data.

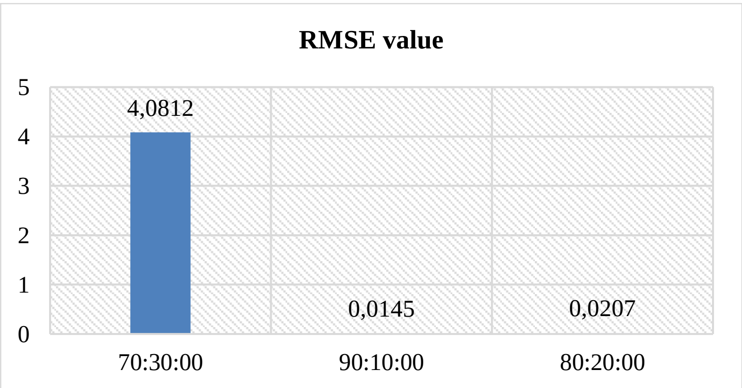

Figure 5. Evaluation of SVM model

In the SVM evaluation using the RMSE value, there are three scenarios that are evaluated with different distribution of training and testing data, namely 70:30, 90:10, and 80:20. In summary, the evaluation of different training and testing splits in this scenario revealed varying RMSE values for the model's predictions. The initial 70:30 split resulted in an RMSE value of 4.0812. However, when the split was adjusted to 90:10, the RMSE value significantly decreased to 0.0145, indicating improved accuracy. Similarly, with an 80:20 split, the RMSE value was slightly higher at 0.0207 but still demonstrated good predictive performance. These findings emphasize the importance of choosing appropriate training and testing splits, as smaller testing sets generally lead to lower RMSE values and improved accuracy in the model's predictions.In the overall evaluation, scenarios with splits of 90:10 and 80:20 show better performance with lower RMSE values. This indicates that the SVM model is able to provide more accurate predictions in this scenario.

3.3 Long Short-Term Memory (LSTM)

The figure 6 shows the prediction results of the Long Short-Term Memory (LSTM) method which are described in 10 years from 2008 to 2018. The test was carried out by measuring the actual Root Mean Square Error (RMSE) value.

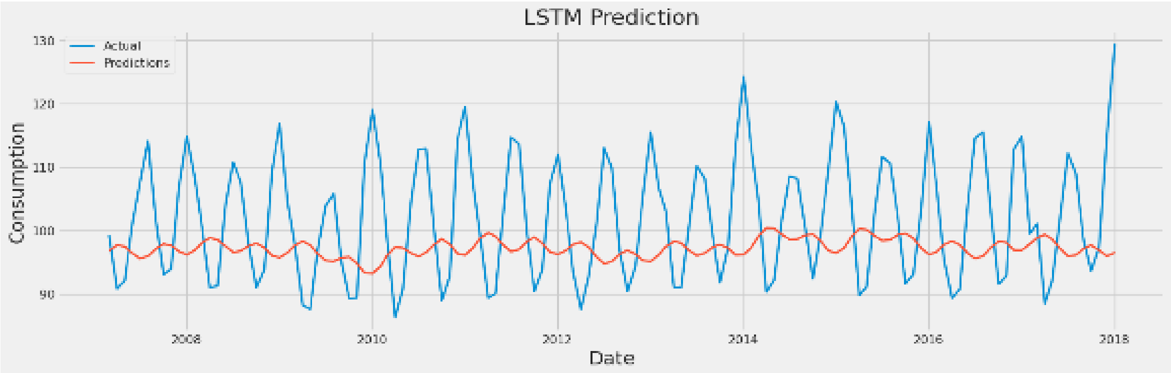

Figure 6. LSTM Prediction Results

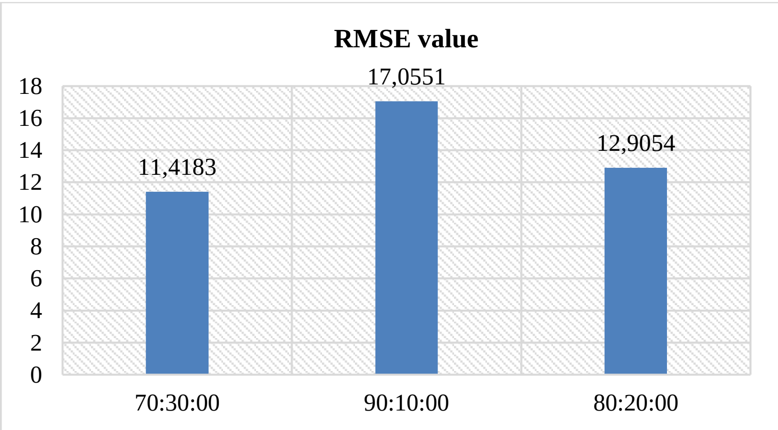

Figure 7. Evaluation of LSTM model

In the LSTM evaluation using the RMSE value, there are three scenarios that are evaluated with different distribution of training and testing data, namely 70:30, 90:10, and 80:20. In summary, the model's performance was evaluated using three different training and testing splits. The 70:30 split yielded the lowest RMSE value of 11.4183, indicating better accuracy compared to the other splits. The 90:10 split resulted in a higher RMSE value of 17.0551, indicating lower accuracy and a larger error. The 80:20 split produced an RMSE value of 12.9054, which was closer to the RMSE value of the 70:30 split.

These findings highlight the importance of choosing an appropriate training and testing split, as it can significantly impact the model's performance. While the 70:30 split performed the best in this analysis, it's essential to consider the specific characteristics and requirements of the dataset when determining the optimal split.In the overall evaluation, all evaluated scenarios show a high RMSE value, which indicates a significant difference between the predicted LSTM value and the actual value. This indicates that the evaluated LSTM model may not provide accurate prediction results in this case.

4. CONCLUSION

Based on the evaluation results of the ARIMA, SVM, and LSTM models in predicting electricity use over the next 1 year, several conclusions can be drawn. ARIMA is able to provide predictions of electricity usage for the next 1 year period. Reliable for several forecasting applications with quite good results. Although ARIMA can provide predictions, the RMSE value is quite high (7.659), indicating a significant deviation between the prediction and the actual value. In this specific context, ARIMA may not be accurate enough to meet the need for high-precision electricity usage predictions.

SVM has a much lower RMSE value (0.020) than ARIMA, indicating a higher level of accuracy in predicting electricity usage. Suitable for applications where high accuracy is a top priority. Even though SVM has a low RMSE value, there are still limitations in its ability to predict electricity usage well over the next 1 year. SVM may require further development or additional methods to improve its performance in this context.

LSTM provides predictions with a lower RMSE value compared to ARIMA (7.659), although it is still higher than SVM (0.020). LSTM can have advantages in handling complex or non-linear patterns in electricity usage data. Although the RMSE value of LSTM is lower than ARIMA, it is still relatively high (11.4183), indicating that LSTM also faces difficulties in predicting accurately in this context. LSTM models may require further parameter adjustments or tuning to improve prediction performance.

In conclusion, although SVM has a lower RMSE value than ARIMA, LSTM shows relatively good ability in predicting electricity usage, especially in dealing with complex patterns. However, the limitations of LSTM are still visible, and it is necessary to carry out further research or parameter experiments to improve its performance or consider using other models that are more suitable to the characteristics of the existing data.

REFERENCES

[1] E. Hossain, MR Tur, S. Padmanaban, S. Ay, and I. Khan, "Analysis and Mitigation of Power Quality Issues in Distributed Generation Systems Using Custom Power Devices," IEEE Access , *vol* . 6, no. c, pp. 16816–16833, 2018, doi: 10.1109/ACCESS.2018.2814981.

[2] NS Syam *et al.* , "Model Support Vector Machine for Predictions on the Use of Electrical Energy in Energy Efficient Homes," *J. Inform.* , vol. 1, no. 2, pp. 56–59, 2022.

[3] S. Rahayu and JJ Purnama, "Classification of Steel Industry Energy Consumption Using Data Mining Techniques," *J. Teknoinfo* , vol. 16, no. 2, p. 395, 2022, doi: 10.33365/jti.v16i2.1984.

[4] M. Danus, "Application of the Moving Average Method for Forecasting the Electrical Load of a 20 Kv Distribution Network at the Simpang Tiga Satu Feeder at the Keramasan Substation," *J. Ampere* , vol. 4, no. 1, p. 252, 2019, doi: 10.31851/ampere.v4i1.2877.

[5] A. González-Briones, G. Hernandez, JM Corchado, S. Omatu, and MS Mohamad, "Machine Learning Models for Electricity Consumption Forecasting: A Review," *2nd Int. Conf. Comput. appl. inf. Secur. ICCAIS 2019* , 2019, doi: 10.1109/CAIS.2019.8769508.

[6] H. Wibowo, Y. Mulyadi, and AG Abdullah, "Forecasting Classified Short Term Electricity Expenses Based on the Autoregressive Integrated Moving Average Method," *Electrans* , vol. 11, no. 2, pp. 44–50, 2012.

[7] A. Arfan and L. ETP, "Comparison of Long Short-Term Memory Algorithm with SVR on Stock Price Prediction in Indonesia," *Petir* , vol. 13, no. 1, pp. 33–43, 2020, doi: 10.33322/petir.v13i1.858.

[8] R. Panggabean, Y. Dewi, and L. Widyasari, "A comparison between Super Vector Regression , Random Forest Regressor , LSTM , and GRU in Forecasting Bitcoin Price," no. November, pp. 17–19, 2022.

[9] AP Point Rahmawati, Landung Sudarmana, "Application of the Arima Box-Jenkins Method for Forecasting Electricity Consumption," *Politeknosains* , vol. XIX, no. 1, pp. 6–11, 2020.

[10] MI Abas and I. Ibrahim, "Optimization of Support Vector Machine Particle Swarm Optimization for Predicting Electrical Energy Consumption," *Jambura J. Informatics* , vol. 1, no. 2, pp. 47–56, 2019, doi: 10.37905/jji.v1i2.2646.

[11] I. Amarulloh, "Forecasting short-term electric power on the Smart Grid Photovoltaic ARIMA method with the effect of temperature in hybrid mode," *J. Tek. Electro* , vol. 10, no. 3, pp. 769–781, 2021.

[12] Y. Chen and K. Wang, "Prediction of satellite time series data based on long short term memory-autoregressive integrated moving average model (LSTM-ARIMA)," *2019 IEEE 4th Int. Conf. Signal Image Process. ICSIP 2019* , pp. 308–312, 2019, doi: 10.1109/SIPROCESS.2019.8868350.

[13] S. Siami-Namini, N. Tavakoli, and AS Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proc. - 17th IEEE Int. Conf. Mach. Learn. appl. ICMLA 2018* , pp. 1394–1401, 2019, doi: 10.1109/ICMLA.2018.00227.

[14] ED Spyrou, I. Tsoulos, and C. Stylios, "Applying and Comparing LSTM and ARIMA to Predict CO Levels for a Time-Series Measurements in a Port Area," Signals , *vol* . 3, no. 2, pp. 235–248, 2022, doi: 10.3390/signals3020015.

[15] Y. Guo, Y. Feng, F. Qu, L. Zhang, B. Yan, and J. Lv, "Prediction of hepatitis E using machine learning models," *PLoS One* , vol. 15, no. September 9, pp. 1–12, 2020, doi: 10.1371/journal.pone.0237750.

[16] A. Hanifa, SA Fauzan, M. Hikal, and MB Ashfiya, "Comparison of LSTM and GRU (RNN) Methods

for Classification of Fake News in Indonesian," Din . *Engineering* , vol. 17, no. 1, pp. 33–40, 2021.
[17] MA Amrustian, W. Widayat, and AM Wirawan, "Evaluation Sentiment Analysis of Teaching Lecturers in Higher Education Using the LSTM Method," *J. Media Inform. Budidarma* , vol. 6, no. 1, p. 535, 2022, doi: 10.30865/mib.v6i1.3527.
[18] RN Yusra, OS Sitompul, and Sawaluddin, "Combination of K-Nearest Neighbor (KNN) and Relief-F to Improve Accuracy in Data Classification," *InfoTekJar J. Nas. inform. and Technol. Jar.* , vol. 1, pp. 0–5, 2021.
[19] Q. Gu, Y. Chang, X. Li, Z. Chang, and Z. Feng, "A novel F-SVM based on FOA for improving SVM performance," *Expert Syst. appl.* , vol. 165, p. 113713, 2021, doi: 10.1016/j.eswa.2020.113713.
[20] M. Mohammadi *et al.* , "A comprehensive survey and taxonomy of the SVM-based intrusion detection systems," *J. Netw. Comput. appl.* , vol. 178, no. January, p. 102983, 2021, doi: 10.1016/j.jnca.2021.102983.
[21] SS Durbha, RL King, and NH Younan, "Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer," *Remote Sens. environment.* , vol. 107, no. 1–2, pp. 348–361, 2007, doi: 10.1016/j.rse.2006.09.031.