Forecasting Climate Change Patterns to Improving Rice Harvest Using SVR for Achieving Green Economy

Carles Juliandy[1], Kelvin[2], Apriyanto Halim[3], Sio Jurnalis Pipin[4], Frans Mikael Sinaga[5], Wulan Sri Lestari[6]
[1,3,4,6]Department of Information Technology, Universitas Mikroskil, Indonesia
[2,5]Department of Informatics Engineering, Universitas Mikroskil, Indonesia
Email: [1]carles.juliandy@mikroskil.ac.id, [2]kelvin.chen@mikroskil.ac.id, [3]apriyanto.halim@mikroskil.ac.id, [4]sio.pipin@mikroskil.ac.id, [5]frans.sinaga@mikroskil.ac.id, [6]wulan.lestari@mikroskil.ac.id

Article Info
Article history:
Received Jul 12th, 2024
Revised Aug 30th, 2024
Accepted Sep 6th, 2024

Keyword:
Climate Change
Climate Forecasting
Green Economy
Support Vector Regression
Rice Harvest

ABSTRACT
The consistently declining rice harvest will cause several economic and environmental problems. The unstable and unpredictable climate change was believed as the main problem of the declining rice harvest. We proposed a method for forecasting climate change to help the farmer in their rice cultivation. We used Support Vector Regression (SVR) to improve algorithm steps such as normalizing the data and applying an Adaptive Linear Combiner (ALC) to optimize the dataset before we processed it with the algorithm. Our model gets 95% accuracy as measured with the confusion matrix. We believe our model will help the farmers in their rice cultivation with good climate forecasting. A further benefit of this research we belief that with the well-forecasted climate, the usage of pesticides will decrease and will help the vision of the Indonesian government with a green economy.

Corresponding Author:
Carles Juliandy,
Departement of Information Technology, Faculty of Informatics
Universitas Mikroskil,
Jl. M.H. Thamrin No. 140, Medan Kota, Kota Medan, Indonesia. 20212
Email: carles.juliandy@mikroskil.ac.id
DOI: http://dx.doi.org/10.24014/ijaidm.v7i2.32393

1. INTRODUCTION

The background of this research was because of consistently declining rice harvest levels year after year. This is evident from the 2023 Dry Unhusked Rice (GKG) yield, which was only 53.98 million tons, the lowest in the last decade, down 33.5% from its peak in 2017 [1], [2]. Based on several previous researches, unstable and unpredictable climate change in many areas of Indonesia was the cause of declining rice harvest levels [3], [4] besides rice diseases and pests [5]. Another effect was the increasing usage of pesticides and other chemical materials to sustain crops during unstable weather. This situation didn't support the green economy concept because the excessive usage of pesticides and chemicals caused environmental damage [6], [7].

The adverse weather if remaining unpredictable will cause several negative effects: 1) Food security will be threatened because Indonesia as a country with the highest rice consumption, will have to import rice if rice production cannot be increased. 2) Unstable price because the demand was higher than supply, as in economic laws it will make the price higher. 3) Unable to achieve a green economy because of high usage of pesticide [7], [8], [9]. The solution to this problem is to use big data to predict climate change based on previous data so that farmers will be better prepared to plant rice, use fewer pesticides, and get better results in rice harvest.

With the big impact of adverse weather on rice harvesting, it's important to forecast the weather change [10], so it can improve the rice harvest. But there are many challenges in forecasting weather such as too much data that can impact the weather forecast and weather data is also the timescale data. So, implementing the machine learning algorithm could help in forecasting the weather [11]

As the grand theory for this research, we found that some researchers conduct supervised learning algorithms in predicting climate for several purposes and also for increasing rice harvest. Research from [12] conducted the use of mLSTM to determine the influence of climate change on rice harvest. Wang [13] applied Support Vector Machine (SVM) for future climate prediction for several purposes. Khatri [14] used AutoRegressive Integrated Moving Average (ARIMA) to predict climate change. The conclusion from the several researches is the algorithm works well in predicting climate change. But, for agriculture, there is not enough to know what will happen with the climate by inputting the real weather conditions. Farmers need to forecast what will happen to the environment in the future for better preparation.

A regression algorithm may work well for predicting future climate based on previous data. Malik [15] states that regression algorithm especially Support Vector Regression (SVR) better for forecasting and SVR also improve from SVM, because it can handle multi-labeled data better also SVM has worse performance if compared to ARIMA and LSTM [16]. For better implementation in real cases such as weather forecasting, the SVR algorithm also has a gap in its implementation. Khosla [17] research combined Artificial Neural Network (ANN) with SVR for increasing crop yield by forecasting climate. The result was not so accurate enough because the data is random and has many labels, as we know the weather data is also random and has many labels.

This research focused on SVR because compared to other algorithms, SVR is more robust to outlier data, and weather data sometimes can be abnormal [18]. SVR is less prone to overfitting, and in climate forecasting, because the overfitting to noise data can lead to poor generalization [19]. SVR is also can perform well with low computer power, so it doesn't need large computational power as compared to other machine learning algorithms such as ANN [20].

SVR has also been implemented in several researches, especially in forecasting. Xu and Jiang, forecast the traffic flow using a deep belief network with SVR and achieve a high prediction result compared to other traditional machine learning models [21]. The research by Quan et.al, [19] focused on forecasting water temperature in reservoirs using SVR. The result was outperforming other machine learning models in accuracy and efficiency metrics. Similar results from Fan et.al, [22] while forecasting short-term electricity load using SVR. The result also gets better accuracy while using SVR compared to another machine learning algorithm.

As a contribution to this research, we also try to improve the algorithm, by adding a step to optimize the data so it can improve the forecasting accuracy. Qu [23] research found labeling the data before processing it will increase the accuracy. Because it makes sure the data has the right target label. So, in this research, we try to combine SVR with the labeling step to forecast future climate change. So, it will help the rice farmer to prepare better for their rice harvesting. Better preparation may also use less pesticide which will support the green economy concept by taking care of the environment

2. RESEARCH METHOD

We proposed a method to increase the accuracy of weather forecasting, so the rice harvest will be increased. Our research method can be seen in Figure 1.

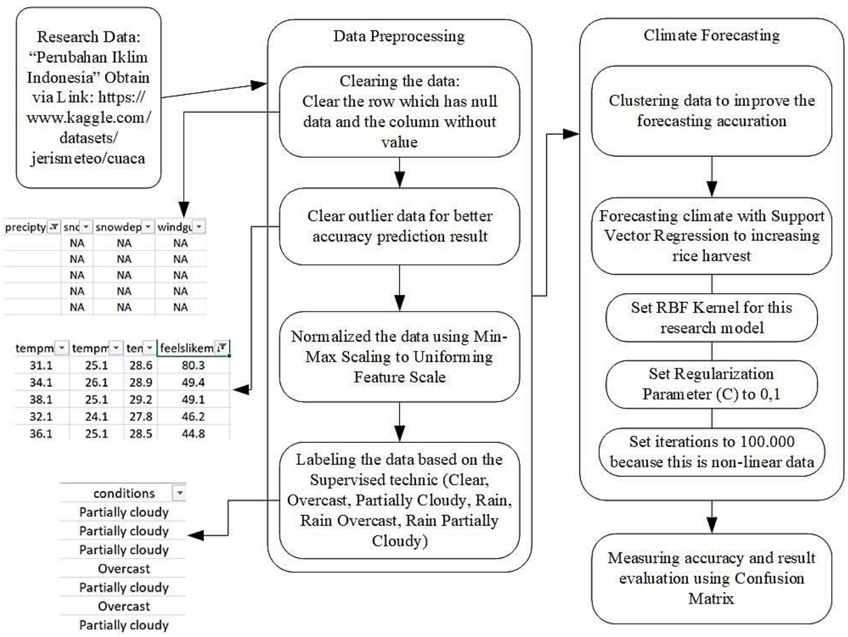

Figure 1. Research Method

Below here will briefly explain the method we used in this research. We used the Indonesia climate data from Badan Metereologi, Klimatologi, dan Geofisika (BMKG) and we got it from Kaggle. The use of Indonesia's climate data gives an accurate image of climate forecasting with the real conditions happening in Indonesia. The data is from 1978-2022, containing 16.391 rows of data and 34 variables. We preprocessed the data by changing the row with a null value into 0 and clearing the column with null data. After clearing the data process there are remaining 25 variables and 15.978 rows data. The cleaned data can be seen in Table 1.

Table 1. Cleaned Data

tempmax	tempmin	temp	feelslikemax	feelslikemin	feelslike	dew	humidity	precip	precipprob	precipcover
32,1	25,1	28,1	39,3	25,1	31,4	24,8	83,5	0,0	0,0	0,0
31,1	25,1	27,9	37,9	25,1	31,8	24,9	84,4	0,0	0,0	0,0
32,1	26,1	29,0	40,8	26,1	33,6	25,5	82,7	0,0	0,0	0,0
33,1	25,1	28,5	39,4	25,1	32,2	25,1	83,2	0,0	0,0	0,0
31,1	24,1	27,1	44,4	24,1	30,7	23,1	85,2	0,0	0,0	0,0
30,1	24,1	26,1	37,6	24,1	27,8	24,7	92,2	0,0	0,0	0,0
31,1	24,1	27,4	42,4	24,1	31,6	25,4	89,7	0,0	0,0	0,0
29,1	24,1	27,0	34,7	24,1	29,6	24,7	87,8	9,0	100	4,17
32,1	25,1	27,0	39,3	25,1	28,7	24,5	87,7	0,0	0,0	0,0

Next step we need to clear the outlier data before we normalized with Min-Max Scaling because the algorithm sensitive to the outlier data [24]. We need to normalize the data because normalized the data can increase the prediction result, especially in data with many units and weather data is so many different units [25]. After clearing the outlier data, the remaining data is 15.969 data, and we also applied the Min-Max algorithm to normalize the data using this formula [24]:

$$X_{\text{scaled}} = (X - X_{min}) / (X_{max} - X_{min}) \tag{1}$$

Where:

X : Original data (value)
X_{min} : Minimum value from the label
X_{max} : Maximum value from the label
X_{scaled} : Scaled data (value)

An example of the applied formula can be seen in Table 2.

Table 2. Normalized Original Data into Scaled Data

Label	Original Data	Scaled Data
Temp	28,1	0,493589
	27,4	0,448717
Feelslike	31,4	0,529680
	27,8	0,365296
dew	24,8	0,906077
	23,1	0,812154
Humidity	83,5	0,710227
	92,2	0,875000

When finished with the normalization, we continued with labeling the data because weather data is combined from multiple model data and we need to make sure the label on the data is for the data and no misplaced label. Based on [26] Adaptive Linear Combiner was ideal for labeling the data. We used these steps of Adaptive Linear Combiner based on [27]:

1. Initialize weights for each data variable $w_1, w_2, ..., w_K$, usually the value is equal (1/K), where K is the total of the variable and Σw_i must be 1. For our data, it has 28 variables, so the weights for each variable are: $w_i = 1/28 = 0{,}0357$
2. For each data, calculate the combined prediction using the formula:

$$\hat{y} = w_1 f_1(x) + w_2 f_2(x) + ... + w_K f_K(x)c \tag{2}$$

Here is an example of combined prediction, which can be seen in Table 3

Table 3. Combined Prediction of Adaptive Linear Combiner

Xscaledtemp	Xscaledfeels	Xscaleddew	Xscaledhumid	ALC
0,493590	0,529680	0,906077	0,710227	0,094232
0,480769	0,547945	0,911602	0,727273	0,095233
0,551282	0,630137	0,944751	0,695076	0,100718
0,519231	0,566210	0,922652	0,704545	0,096841
0,429487	0,497717	0,812155	0,742424	0,088599
0,365385	0,365297	0,900552	0,875000	0,089472
0,448718	0,538813	0,939227	0,827652	0,098332

After the combined prediction, we need to calculate the error of the combined prediction and also for each variable error. The formula to count of each variable error is:

$$e_i = |y - f_i(x)| \text{ for } i = 1 \text{ to } K \tag{3}$$

where the formula to count the error of combined prediction is:

$$e\_combined = |y - \hat{y}|e \tag{4}$$

After we get the error value, we need to update the weight with the formula:

$$w'_i = w_i * \beta\text{^}|y - f_i(x)| \tag{5}$$

3. Repeat step 2 for all instances in the training set until the error value is small enough, and in this research, our threshold is 0,005.

When the data is already labeled, before we can process it with the SVR algorithm we convert the target variables to the integer type. So, we can measure the accuracy. Because the target variable was string type. In the data there are 6 types of targets and here the target and converted result

1. Overcast → 0
2. Partially Cloudy → 1
3. Clear → 2
4. Rain, Overcast → 3
5. Rain, Partially Cloudy → 4
6. Partially Cloudy → 5

For the SVR algorithm, we used the steps based on [28]. The details of each step can be seen below:

1. With the training dataset form $((x_1, y_1), ..., (x_n, y_n))$ where x_i was the input vector and y_i was the target value, we can use the simple SVR function that can be expressed as:
2.

$$f(x) = w * \emptyset(x) + b \tag{6}$$

Where:

w : weight vector in the feature space
$\emptyset$: transfer function
b : bias

3. Find the optimal SVR function $f(x)$ to minimize $||w||^2$ and to ensure all residual in ε was insensitive. The optimized formulation can be expressed as:

$$Minimize\ \frac{1}{2}\ ||w||^2 + C\ \sum_{i=1}^{N} \xi_i + \ \xi_i^* \tag{7}$$

$$\text{Subject to: } \begin{cases} Y_i - f(x) \leq \varepsilon + \xi_i \\ f(x) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, 3, \ldots, N \end{cases} \tag{8}$$

Where ξ_i and ξ_i^* are the slack and the $-\varepsilon$ and $+\varepsilon$ are the distance between the hyperplane and the boundary lines. Because of the limitation of a high-speed computing system, linear support vector regression can be expressed as:

$$y = \sum_{n=1}^{N} (\propto_n - \propto_n^*) K(x_n . x) + b \tag{9}$$

Where, $\propto_n, \propto_n^*$ and $K(x_n \, . \, x)$ are dual variables and kernel function, which is in this research we used radial basis function (RBF) kernel function because it can give better accuracy [25].

4. The RBF kernel function formula can be expressed as:

$$K(x, x_i) \; = \; exp(-\gamma \| x_i \; - \; x^2 \|) \tag{10}$$

Where, γ was the kernel parameter, which means that C, γ, and ε are the parameters that were responsible for the SVR performance.

After the process, we measure the ability of our model using the confusion matrix [29], the rule used in this research can be seen below:
1. True Positive (TP) → When the predicted result and the actual result are the same.
2. False Positive (FP) → When the predicted result is different from the actual result.
3. False Negative (FN) → When the predicted result is different from the actual result.

We will evaluate several metrics from the confusion matrix based on previous research [30] such as:
1. Precision, with the formula

$$Precision \; = \frac{TP}{TP + FP} \tag{11}$$

2. Recall, with the formula

$$Recall \; = \frac{TP}{TP + FN} \tag{12}$$

3. F1-Score, with the formula

$$F1 \; = \; 2 \, x \, \frac{Precision \; x \; Recall}{Precision + Recall} \tag{13}$$

4. Support for each class, is how many actual examples are included in the class
5. Accuracy, with the formula

$$Accuracy \; = \frac{\sum TP}{\sum (TP + FP + FN)} \tag{14}$$

3. RESULTS AND ANALYSIS

The model works well in predicting climate change based on previous data. The prediction result of our model can be seen in Table 4.

Table 4. Prediction Result

Prediction Result	Predicted Label
2	Rain, Partially cloudy
2	Rain, Partially cloudy
2	Rain, Overcast
1	Rain, Partially cloudy
2	Rain, Partially cloudy
1	Partially cloudy
2	Rain, Partially cloudy
5	Rain, Partially cloudy
2	Overcast
2	Rain, Partially cloudy

With 80% of training data and 20% of testing data, we compare the predicted data with the actual data to measure our model capability. We compare each data and find if it is the same as the actual data or not. The sample of the comparison result on the predicted result with the actual data can be seen in Table 5.

Table 5. Comparison of Prediction Data with The Actual Data

Actual Data	Actual Label	Predicted Result	Predicted Label
5	Rain, Partially cloudy	5	Rain, Partially cloudy
5	Rain, Partially cloudy	5	Rain, Partially cloudy

Actual Data	Actual Label	Predicted Result	Predicted Label
4	Rain, Overcast	4	Rain, Overcast
5	Rain, Partially cloudy	5	Rain, Partially cloudy
5	Rain, Partially cloudy	5	Rain, Partially cloudy
2	Partially cloudy	2	Partially cloudy
5	Rain, Partially cloudy	5	Rain, Partially cloudy
5	Rain, Partially cloudy	5	Rain, Partially cloudy
1	Overcast	1	Overcast
5	Rain, Partially cloudy	5	Rain, Partially cloudy

From the comparison result our model can gain 95% of accuracy on the confusion matrix. For the detail of the comparison result, Figure 2 on the next page will show the result.

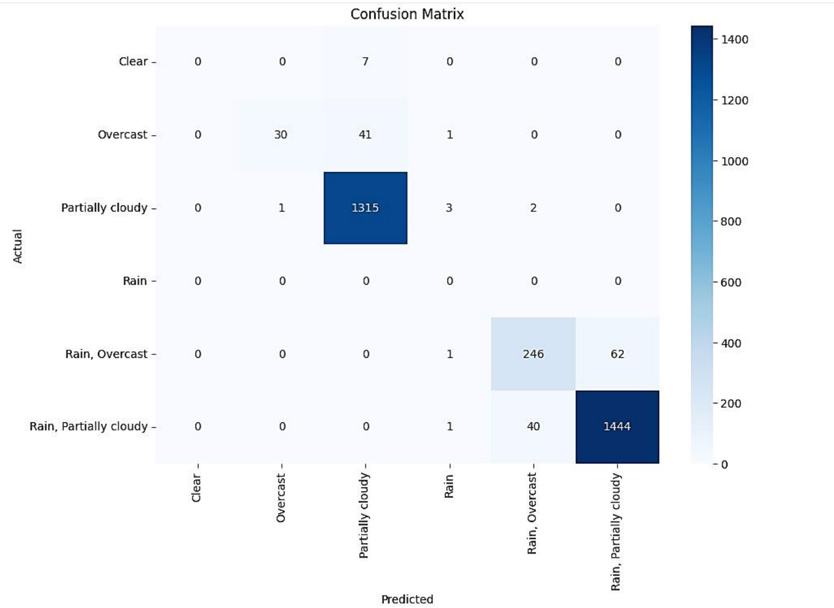

Figure 2. Actual Result VS Prediction Result

From Figure 2, we know that there are a lot of predicted results categorized as True Positive. The most actual labels in the testing data were the Partially Cloudy and the Rain, Partially Cloudy labels. What we can explain is, that in the testing data, the data with Partially cloudy label data have about 1.321 actual data. But in the prediction result, there are about 1.315 predicted with Partially Cloudy, 1 predicted with Overcast, 3 predicted with Rain, and 2 predicted with Rain, Overcast. It means that in Partially Cloudy label there is 99,5% accuracy.

Similar result with Rain, Partially Cloudy. It gets 97,2% accuracy from 1.444 true positives with 1.485 total predictions. There are 40 wrong predicted with Rain, Overcast, and 1 wrong predicted with Rain. The worst result was the Overcast label with an accuracy of 41,67%. This percentage result was obtained by dividing 30 true positive results with 72 all prediction data. This result may be understandable because Indonesia is a tropical country and the climate is majority with high humidity weather such as partially cloudy, rainy, partially cloudy, or rainy, overcast [31]. The more detailed testing result can be seen in table 6.

Table 6. Testing Result Each Label and Evaluation

	Precision	Recall	F1-Score	Support
Clear	0.00	0.00	0.00	7
Overcast	0.97	0.42	0.58	72
Partially Cloudy	0.96	1.00	0.98	1321
Rain	0.00	0.00	0.00	0
Rain, Overcast	0.85	0.80	0.82	309
Rain, Partially Cloudy	0.96	0.97	0.97	1485
Accuracy			0.95	3194
Macro Avg.	0.62	0.53	0.56	3194
Weighted Avg.	0.95	0.95	0.95	3194

The testing result of each label showed that the average accuracy of our model was 0,95 or 95% from 3.194 predicted data. The average precision from the model was 0,62 or 62% which means that 62% of true positive prediction was exactly true positive[32]. This result was caused by the clear label and the rain label did not have a true positive result. We assume that this is because the dataset has less clear and rain-label data. It was also supported by the dataset which is only 54 from 15.978 data with the rain and clear label.

The average recall from our model is 0,53 or 53%, this result aligned with the precision because recall is also impacted by the true positive metric. The f1-score was the harmonics average between recall and precision. We got 0,56 for our average f1-score which is balanced enough for recall and precision value.

From the result, we believe that our model was capable of forecasting climate change in the future based on the previous data. The 95% accuracy was high enough because the weather data was combined with the multiple model data. The confusion matrix rule can be seen in Table 6.

Table 6. Confusion Matrix Result

		Predicted Result	
		Positive	Negative
Actual Result	Positive	TP 3035	FN -
	Negative	FP 159	TN -

From Table 6, we can conclude that the correct predicted result was 3.035 of 3.194 testing data. Which is about 95% of accuracy. The correct predicted result means if the actual label was clear and the predicted label was also clear and so on. The wrong predicted result was 159 data. In this research, there are no False Negative and True Negative because the data is multi-labeled so there is only true or false prediction.

4. CONCLUSION

The consistently declining rice harvest was caused by several problems. One of the main problems was the adverse weather was not well predicted. We implemented Support Vector Regression (SVR) with several improvements in the SVR algorithm such as normalization and Adaptive Linear Combiner (ALC) to increase the accuracy of forecasting, because weather data have so many different units. Our model has 95% accuracy as measured by the confusion matrix. We believed that our model was capable of helping farmers forecast climate change in the future to help the farmer increase their rice harvest. The well-forecasted climate also helps farmers in reducing the usage of pesticides. So, less usage of pesticides will be more environmentally friendly, which aligns with the Indonesian government's vision of a green economy. Future research can be done in several ways, such as applying this method with other climate data with a more capable amount of lable. The improvement can also be done by combining the algorithm with the other algorithm to increase the accuracy by nearly 100%.

ACKNOWLEDGEMENTS

This research was supported by Direktorat Jendral Pendidikan Tinggi, Riset dan Teknologi, and also supported by another author as mentioned in the author list. Also special thanks to Universitas Mikroskil for giving us much training, so we could improve our research skills and we were accepted as the grant recipient of Penelitian Dosen Pemula (PDP) 2024.

REFERENCES

[1] C. M. Annur, “Produksi Padi Indonesia Cenderung Menurun dalam 10 Tahun Terakhir.”
[2] Badan Pusat Statistik, “Luas Panen dan Produksi Padi di Indonesia 2023,” 2023. Accessed: Mar. 26, 2024. [Online]. Available: https://www.bps.go.id/id/pressrelease/2024/03/01/2375/pada-2023--luas-panen-padi-mencapai-sekitar-10-21-juta-hektare-dengan-produksi-padi-sebesar-53-98-juta-ton-gabah-kering-giling--gkg-.html
[3] M. L. Hidayatullah and B. U. Aulia, “Identifikasi Dampak Perubahan Iklim Terhadap Produksi Padi,” *Jurnal Teknik ITS*, vol. 8, no. 2, pp. 43–48, 2019, doi: 10.12962/j23373539.v8i2.49241.
[4] A. B. Santoso, T. Supriana, and M. A. Girsang, “Pengaruh Curah Hujan pada Produksi Padi Gogo di Indonesia,” *Jurnal Ilmu Pertanian Indonesia*, vol. 27, no. 4, pp. 606–613, Oct. 2022, doi: 10.18343/jipi.27.4.606.
[5] D. Hastari, S. Winanda, A. R. Pratama, N. Nurhaliza, and E. S. Ginting, “Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease,” *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 2, Feb. 2024, doi: 10.57152/predatecs.v1i2.865.
[6] Makmun, “Green Economy: Kosep, Implementasi, dan Peranan Kementrian Keuangan”.
[7] S. Skendžić, M. Zovko, I. P. Živković, V. Lešić, and D. Lemić, “Review: The impact of climate change on agricultural insect pests,” *Insects*, vol. 12, no. 5, May 2021, doi: 10.3390/insects12050440.
[8] T. Mulyaqin, “The Impact of El Niño and La Nina on Fluctuation of Rice Production in Banten Province,” *Agromet*, vol. 34, no. 1, pp. 34–41, May 2020, doi: 10.29244/j.agromet.34.1.34-41.
[9] G. S. Malhi, M. Kaur, and P. Kaushik, “Impact of climate change on agriculture and its mitigation strategies: A review,” *Sustainability (Switzerland)*, vol. 13, no. 3, pp. 1–21, Feb. 2021, doi: 10.3390/su13031318.
[10] N. T. Luchia, E. Tasia, I. Ramadhani, A. Rahmadeyan, and R. Zahra, “Performance Comparison Between Artificial Neural Network, Recurrent Neural Network and Long Short-Term Memory for Prediction of Extreme Climate

Change," *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 2, pp. 62–70, Feb. 2024, doi: 10.57152/predatecs.v1i2.864.

[11] M. Chantry, H. Christensen, P. Dueben, and T. Palmer, "Opportunities and challenges for machine learning in weather and climate modeling: Hard, medium and soft AI," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, Apr. 2021, doi: 10.1098/rsta.2020.0083.

[12] S. Thirumal and R. Latha, "Machine Learning based Predictive Assessments of Impacts of Influential Climatic Conditions for the Sustainable Productivity of Paddy Crops," *SSRG International Journal of Electronics and Communication Engineering*, vol. 10, no. 5, pp. 116–127, May 2023, doi: 10.14445/23488549/IJECE-V10I5P111.

[13] Y. Wang, X. Wang, X. Li, W. Liu, and Y. Yang, "Future Climate Prediction Based on Support Vector Machine Optimization in Tianjin, China," *Atmosphere (Basel)*, vol. 14, no. 8, Aug. 2023, doi: 10.3390/atmos14081235.

[14] P. Khatri, T. Arjariya, and N. S. Mitra, "Climate change forecasting using data mining algorithms," *Aqua Water Infrastructure, Ecosystems and Society*, vol. 72, no. 6, pp. 1065–1083, Jun. 2023, doi: 10.2166/aqua.2023.046.

[15] A. Malik, Y. Tikhamarine, D. Souag-Gamane, O. Kisi, and Q. B. Pham, "Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction," *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 11, pp. 1755–1773, Nov. 2020, doi: 10.1007/s00477-020-01874-1.

[16] N. W. Azani, C. P. Trisya, L. M. Sari, H. Handayani, and M. R. M. Alhamid, "Performance Comparison of ARIMA, LSTM and SVM Models for Electric Energy Consumption Analysis," *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 2, Feb. 2024, doi: 10.57152/predatecs.v1i2.869.

[17] E. Khosla, R. Dharavath, and R. Priya, "Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression," *Environ Dev Sustain*, vol. 22, no. 6, pp. 5687–5708, Aug. 2020, doi: 10.1007/s10668-019-00445-x.

[18] Z. Zhang and W. C. Hong, "Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads," *Knowl Based Syst*, vol. 228, Sep. 2021, doi: 10.1016/j.knosys.2021.107297.

[19] Q. Quan, Z. Hao, H. Xifeng, and L. Jingchun, "Research on water temperature prediction based on improved support vector regression," *Neural Comput Appl*, vol. 34, no. 11, pp. 8501–8510, Jun. 2022, doi: 10.1007/s00521-020-04836-4.

[20] R. Rodríguez-Pérez and J. Bajorath, "Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery," *J Comput Aided Mol Des*, vol. 36, no. 5, pp. 355–362, May 2022, doi: 10.1007/s10822-022-00442-9.

[21] H. Xu and C. Jiang, "Deep belief network-based support vector regression method for traffic flow forecasting," *Neural Comput Appl*, vol. 32, no. 7, pp. 2027–2036, Apr. 2020, doi: 10.1007/s00521-019-04339-x.

[22] G. F. Fan, M. Yu, S. Q. Dong, Y. H. Yeh, and W. C. Hong, "Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling," *Util Policy*, vol. 73, Dec. 2021, doi: 10.1016/j.jup.2021.101294.

[23] X. Qu *et al.*, "A Survey on the Development of Self-Organizing Maps for Unsupervised Intrusion Detection".

[24] S. Sinsomboonthong, "Performance Comparison of New Adjusted Min-Max with Decimal Scaling and Statistical Column Normalization Methods for Artificial Neural Network Classification," *Int J Math Math Sci*, vol. 2022, 2022, doi: 10.1155/2022/3584406.

[25] A. Ambarwari, Q. J. Adrian, and Y. Herdiyeni, "Analisis Pengaruh Data Scaling Terhadap Performa Algoritme Machine Learning untuk Identifikasi Tanaman," *Jurnal RESTI*, vol. 1, no. 3, pp. 117–122, 2020.

[26] B. Han and B. Bae, "Novel phase-locked loop using adaptive linear combiner," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 513–514, Jan. 2006, doi: 10.1109/TPWRD.2005.860436.

[27] Y. Singh, I. Hussain, B. Singh, and S. Mishra, "Single-phase solar grid-interfaced system with active filtering using adaptive linear combiner filter-based control scheme," *IET Generation, Transmission and Distribution*, vol. 11, no. 8, pp. 1976–1984, Jun. 2017, doi: 10.1049/iet-gtd.2016.1392.

[28] R. Taghizadeh-Mehrjardi *et al.*, "Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models," *Geoderma*, vol. 383, Feb. 2021, doi: 10.1016/j.geoderma.2020.114793.

[29] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf Sci (N Y)*, vol. 507, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.

[30] F. Mikael Sinaga, P. Sirait, and A. Halim, "Optimization of SV-kNNC using Silhouette Coefficient and LMKNN for Stock Price Prediction," in *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 326–331. doi: 10.1109/ISRITI51436.2020.9315516.

[31] K. Bali and S. Maggu, "Weather Forecasting Using KNN in Machine Learning," *International Journal of Advances in Engineering and Management (IJAEM)*, vol. 3, no. 6, pp. 2395–5252, 2021, doi: 10.35629/5252-0306249255.

[32] M. Yunus, M. K. Biddinika, and A. Fadlil, "Classification of Stunting in Children Using the C4.5 Algorithm," *Jurnal Online Informatika*, vol. 8, no. 1, pp. 99–106, Jun. 2023, doi: 10.15575/join.v8i1.1062.

BIBLIOGRAPHY OF AUTHOR

Carles Juliandy, S.Kom., M.Kom., The author is a lecturer at the Department of Information Technology, Faculty of Informatics, Universitas Mikroskil, Medan. He completed a Bachelor's degree in Informatics Engineering at STMIK TIME and a Master's degree in Informatics Technology at STMIK Mikroskil. The author completed several research, especially in Blockchain technology and computer vision.

Kelvin, S.Kom., M.Kom., The author is a Software Engineer and Lecturer in Informatics Engineering, at the Faculty of Informatics, Mikroskil University, Medan. He completed his bachelor's degree in Informatics Engineering at STMIK Mikroskil in 2018. Then, in 2020, the author pursued postgraduate studies in Information Technology at Mikroskil University and completed them in 2021. The courses he has taught include Introduction to Algorithms, Web Design, C Programming, Object-Oriented Programming, Back-End Web Development, Artificial Intelligence, Natural Language Processing, and Machine Learning.

Apriyanto Halim, S.Kom., M.Kom., A researcher and permanent lecturer in the Information Technology study program, Faculty of Informatics at Mikroskil University, Medan. Researchers often conduct research in the field of image processing that focuses on image quality. This can be used for further processes in determining patterns or for the process of detecting garbage. The research currently being conducted focuses on the process of detecting the seabed using YOLO.

Sio Jurnalis Pipin, S.Kom., M.Kom., Lecturer in Department of Technology Information, Faculty of Informatics, at Universitas Mikroskil, Medan. His educational background includes a Bachelor of Computer Science and a Master of Information Technology, both earned from Universitas Mikroskil. His research interests primarily focus on Computer Science, Artificial Intelligence, and the general field of Information Technology.

Frans Mikael Sinaga, S.Kom., M.Kom., Lecturer at the Department of Informatics Engineering, Faculty of Informatics, Mikroskil University, Medan. Born in Penggalangan village on October 24, 1993. The author is the third child out of 4 siblings of Mr. Waristo and Mrs. Linda. The author completed a Bachelor's degree (S1) in Informatics Engineering and a Master's degree (S2) in Information Technology at STMIK Mikroskil Medan. The author has written several book titles such as Introduction to Computer Networks and Data Mining. In addition to writing books, the author has also conducted several research projects in the fields of Data Science and Computer Vision.

Wulan Sri Lestari, S.Kom., M.Kom., The author is a permanent lecturer in the Information Technology Study Program at the Faculty of Informatics, Universitas Mikroskil, Medan. The author completed a Bachelor's degree in Informatics Engineering and a Master's degree in Information Technology at STMIK Mikroskil Medan. The author has written several books, including Introduction to Databases and Fundamentals of Programming. In addition to writing books, the author has also conducted research in the fields of Cryptography, Data Mining, and Artificial Intelligence.