International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Flood Prediction Based on Weather and Water Level Historical Data Using Recurrent Neural Networks: A Case Study of Jakarta Flood Incidents

Hanis Amalia Saputri [1], Diaz D Santika [2]

Submitted: 10/09/2022 **Accepted:** 20/12/2022

Abstract: In spite of high sea tides in combination with subsidence cause floods in the northern part of the city, successive rainstorm along the year is considered as the decisive cause of flood incidents in the Jakarta area. Flood incidents can massively damage and inevitably disrupt most of social-economic activities of the city. Taking flood risk in many respects into account, a local flood early warning services (FEWS) as the integral part of the city flood comprehensive mitigation plan may be urgently needed. The capability of FEWS to provide a prediction of the scale, timing, and location of the impending flood may then be used to take city-wide precautionary steps. In this study, based on local weather and floodgate water level historical data, an attempt to develop a base model of such a FEWS using recurrent neural networks (RNN) is carried out. The local weather time series data is first concatenated with the floodgate water level data and it is then utilized to predict water level at the corresponding floodgate in 7 days ahead. The predicted water levels in turn are used to decide flood alert categories in the nearby areas surrounding the floodgates. Different types of RNN such as long short-term memory (LSTM), gated recurrent unit (GRU) and combination of LSTM-GRU are examined in order to get the best model capable of giving minimum prediction error. Our computational experiments show that all three models succeed to produce a considerable low prediction error on three different datasets with the stacked GRU model demonstrates its superiority compared to the other two models. The stacked GRU with 32 neuron each is capable of giving root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) namely 109.73, 82.91, and 0.05 respectively on Marina Ancol floodgate dataset

Keywords: Early warning system, flood prediction, gated recurrent unit, long short-term memory, recurrent neural networks.

1. Introduction

Jakarta is a special province in Indonesia with 662 km2 of low, flat basin that has an average elevation of 8m above the sea level with 40% of the area, particularly northern areas, is below mean sea level (MSL). Based on a study conducted by Garschagen, et al [1], Jakarta is one of the cities with the highest flood risk in the world. In spite of land subsidence and sea level rise cause floods in the northern part of the city, extreme rainfall and weather changes along the year constitute a decisive cause of flood incidents in the Jakarta area. Based on the analysis of rainfall and air temperature data collected from the year 1991-2020, Arsyad, M [2] estimates that Indonesia's rainfall rate will continually increase by 8.2mm/year along with an increase of maximum air temperature of 0.0317^{o}C/year. In addition to these weather issues, increasing population pressure and subsidence (10 cm/year or more) of areas already below MSL lead to an autonomous increase of flood risk. Taking flood risk in many respects into account, flood early warning services (FEWS) as the integral part of the city flood comprehensive mitigation plan may be urgently needed.

[1] *Computer Science Department, Binus Graduate Program, Master of Computer of Science, Bina Nusantara University, Jakarta 11480, Indonesia ORCID ID : 0000-0002-4068-9653*
[2] *Computer Science Department, Binus Graduate Program, Master of Computer of Science, Bina Nusantara University, Jakarta 11480, Indonesia ORCID ID : 0000-0002-9971-0744*
** Corresponding Author Email: hanis.saputri@binus.ac.id*

Since both weather data from meteorological stations and water level historical data from hydrological stations typically have a time series structure, most of the researchers recently prefer to use sequence-to-sequence deep learning-based models to forecast future data. As one of deep learning-based approaches, recurrent neural networks (RNN) has been a notorious solution to various types of forecasting problems. Study by Faruq, et al [3] examined and compared LSTM as the state-of-the-art RNN with the radial basis function neural network (RBFNN) for water level prediction. The research intend to forecasts river water level based on a single time series data with time steps and target values correspond to hourly data and water level respectively. They succeeded to get 0.20593 RMSE and coefficient of determination (R^2) as high as 0.844. Based on their research results, it is concluded that LSTM networks is a promising alternative technique to the solution of flood modelling and forecasting problems.

Another related study by Chhetri, et al [4], proposed an approach to rainfall prediction by combining bidirectional long short-term memory (BLSTM) and gated recurrent unit (GRU). This study compares the proposed approach with six other machine learning based models namely linear regression, multi-layer perceptron (MLP), convolutional neural networks (CNN), LSTM, GRU, and BLSTM. The experiment shows that BLSTM-GRU gives the best prediction accuracy with MSE score of 0.007. Chu, et al [5] made an attempt to observe the performance of LSTM, GRU and the combination of LSTM and GRU layers. Employing three different actual observation data, it is found that the LSTM-GRU model achieve best results on most datasets, while the LSTM model get the best prediction accuracy only on one of them.

In this study, an attempt to build a base model of the Jakarta FEWS using RNN is carried out. The Jakarta FEWS preliminary model employing local weather and floodgate water level historical data is designed to have the capability of predicting flood in 7 days ahead. Different types of RNN such as stacked LSTM, stacked GRU and combination of LSTM-GRU are examined to find the best FEWS preliminary model capable of giving the lowest prediction error. Since the local weather data and floodgate water level data are taken from different sources and available in different interval basis, in this study an extra effort has also been performed to build an integrated time series data by preprocessing and concatenating local weather and floodgate water level historical data collected from the year 2017-2020.

2. Methodology

As we intend to predict water level at the floodgates based on the weather variables, the local weather time series data must also be provided. Figure 1 shows the overall procedures to perform such a prediction, started from concatenating water level and local weather time series data to be an integrated time series datasets so that the relation between weather variables and water level is easily examined. After replacing missing values, removing outliers, converting categorical data types into float numerical data types, and data normalization, the concatenated time series dataset is then transformed into supervised learning type dataset containing not only sequences of variables input but also water level data as the target variables.

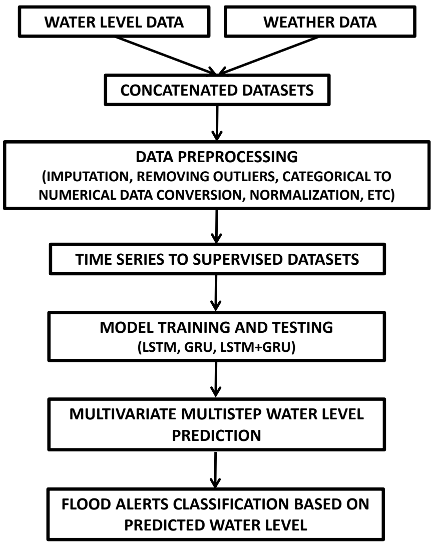

Fig 1. Methodology

The water level time series data from various local floodgates is collected from Open Data Jakarta and it is available from the year 2017-2020 only. The dataset, available in 30 minutes interval, mainly comprises of date, time, name of floodgate, location, the scale of water level (mm), and flood alert categories. Following the availability of the water level dataset, the Jakarta local weather data from the year 2017-2020 was also collected from the Indonesia's Meteorology, Climatology, and Geophysics Agency (BMKG). The local weather dataset available in daily interval contains related information namely minimum temperature (Tn), maximum temperature (Tx), average temperatures (Tavg), average humidity (RH_avg), rainfall rate(RR), duration of solar radiation (ss), maximum wind speed (ff_x), wind direction at maximum speed (ddd_x), average wind speed (ff_avg), and wind direction most (ddd_car). As the interval of weather dataset is daily basis, prior to concatenating both datasets the water level dataset is first down-sampled from hourly basis into daily interval. Figure 2 shows partially the concatenated dataset from which the floodgate water level in 7 days ahead will be predicted. It can be observed from Figure 2, our study is going to deal with 10 weather variables as the multivariate time series input and the scale of water level (mm) as the target vector output. The predicted water level in turn will be used to classify flood alerts categories. In this study, we focus only on several crucial floodgates prone to cause serious flood incidents in the local area surrounding the floodgate. They are floodgates Pasar Ikan, Marina Ancol, and Karet. The total length of each concatenated dataset collected from the year 2017-2020 is 1461 days with 11 features including the water level variable. Each dataset is then splitted up for training and validation.

Fig 2. Partial concatenated dataset

	Tn	Tx	Tavg	RH_avg	RR	ss	ff_x	ddd_x	ff_avg	ddd_car	water_level
date											
2017-01-01	26.0	34.4	29.9	70	0.20	5.0	4	300	2	W	1843.768116
2017-01-02	27.0	34.2	30.2	68	0.00	7.5	4	290	2	N	1843.768116
2017-01-03	26.0	33.8	29.9	71	0.00	8.2	6	310	2	N	1868.045113
2017-01-04	25.0	30.4	27.4	79	38.70	5.0	3	250	2	NW	1884.485294
2017-01-05	25.0	31.4	28.0	81	9.00	5.0	4	270	1	N	1791.205674
...	...	...	...	...	...	...	...	...	...	...	...
2020-12-27	25.6	32.4	29.4	70	0.50	4.8	6	330	3	NW	1825.652174
2020-12-28	26.4	33.0	29.1	70	0.25	7.0	4	310	3	W	1702.272727
2020-12-29	26.4	32.0	28.3	73	0.00	5.3	9	290	2	C	1759.090909
2020-12-30	25.6	32.6	28.0	75	2.50	1.5	5	270	2	W	1648.214286
2020-12-31	25.6	30.2	26.5	83	5.00	0.2	4	240	1	C	1727.837838

1461 rows × 11 columns

The model development is carried out by reshaping 2D supervised dataset into 3D input array compatible with the model input, and followed by training and testing three powerful RNN models namely stacked LSTM, stacked GRU, and stacked LSTM + GRU. Since the model is expected to have the capability of predicting flood in seven days or a week ahead, in this study we apply "walk forward validation" strategy with vector output of length 7. Through this strategy the actual data is made available to the model so that it can be used as the basis for making a prediction on the subsequent week. The "walk forward validation" comprises of moving along the time series one-time step at a time, for example data sequence from week1 is used to predict week2, and data sequence from week1 and week2 is then utilized to predict week3, and so forth. The very last vector output that contains water level prediction in 7 days ahead is eventually obtained and used to classify the flood alert categories.

3. Recurrent Neural Networks

As it is shown in Figure 3, recurrent neural networks (RNN) is a type of artificial neural network that has a conceptual delay block through which h^{t-1} is allowed to feed back into the hidden layer. With such an architecture, the RNN will be able to retain information in the earlier parts of data sample in the memory and move it to the later parts of the same data sample to ensure better knowledge discovery [6]. However, in practice a simple RNN loss its capability in handling steps with long sequences as it encounters the vanishing gradient problem during training [7]. The gradient of the loss function of the network gets extremely high in the presence

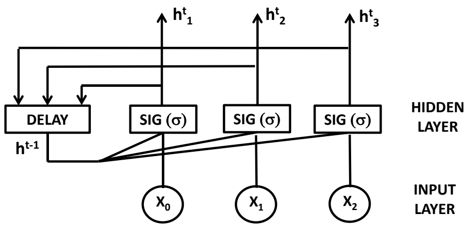

Fig 3. Simple RNN architecture

of such a vanishing gradient problem and in consequence may degrade the accuracy of the trained networks. More complex RNN such as LSTM or GRU networks may be needed to overcome the aforementioned problem. LSTM or GRU networks having gate functions within its structure will be able to control what information will be passed to the memory cell based on previous output and current sensor measurement data, how the memory cell will be updated, and how to control which information will be carried to next time-step (see Figure 4).

3.1. Long Short-Term Memory

Each unit of LSTM networks consisting of 3 gate functions and cell states can carries information from the beginning up till the end part of the time-steps without worrying to get vanished. The first gate function, the forget gate, will decide information from previous hidden state h_{t-1} and input X_{t-1} ether must be ignored or to be further processed to the next state. As it can be observed from Figure 4, such a "go-no go" decision in the forget gate is implemented using sigmoid activation function. The next gate function that is the input gate is responsible to controll what relevant information must be entered and added to the existing information by updating the cell states. The output gate as the last gate function is in charge to control which information will be

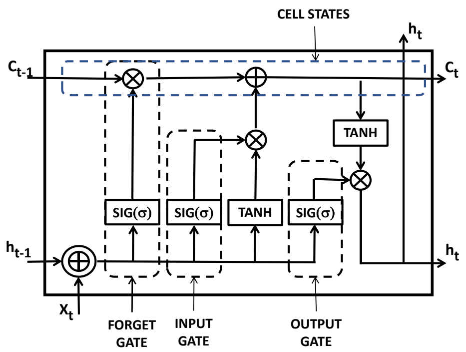

Fig 4. LSTM Architecture

conveyed to the next time-steps.

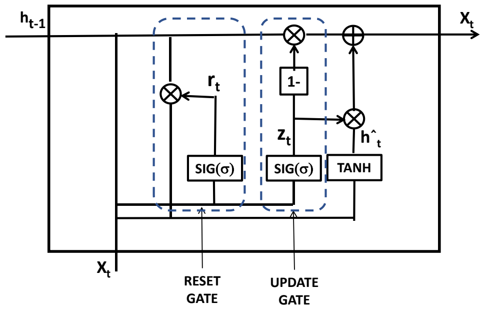

Fig 5. GRU Architecture

3.2. Gated Recurrent Unit

As it is illustrated in Figure 5, GRU networks, the newer generation of RNN, has two gates namely reset and update gates within each unit of the networks and use the hidden state to transfer information. The update gate yields Z_t at time step t using formula $Z_t = \sigma(W^{(z)}X_t + U^{(z)}h_{t-1})$. When X_t is plugged into the network unit, it is multiplied by its own weight $W^{(z)}$. The same goes for h_{t-1} which holds the information for the previous $t-1$ units and is multiplied by its own weight $U^{(z)}$. Both results are added together and a sigmoid activation function is applied to squash the result between 0 and 1. This is to mention that the update gate helps the model to determine how much of the past information including the current one needs to be passed along to the future. The model can decide to copy all the information from the past and eliminate the risk of vanishing gradient problem. The other gate, the reset gate, essentially is used from the model to decide how much of the past information to forget. It uses a very similar formula as the one for the update gate, the difference comes in the weights and the gate's usage. If carefully trained, both LSTM and GRU can perform extremely well even in complex scenarios.

4. Result and Analysis

The models are designed to receive 10 weather features wrapped up in past 7 days sliding window or more and to yield a vector output of length 7 as we expect they could be able to predict water level in 7 days ahead. Each model namely stacked LSTM-LSTM, stacked GRU-GRU, and stacked LST-GRU is trained and validated using three different concatenated datasets. Employing similar hyperparameters setting as is shown in Table 1 for each model, we found that all three models succeed to obtain a considerable low mean squared error (MSE) on three different datasets. The stacked GRU-GRU model with 32 neurons each, however, is capable of giving a slightly lower prediction error compared to the other two models especially when the model is applied to Marina Ancol dataset. Figure 6 shows training and validation MSE losses of the stacked GRU model on three different datasets. It is easily observed from Figure 6 that the stacked GRU model has better training and validation losses on Marina Ancol dataset.

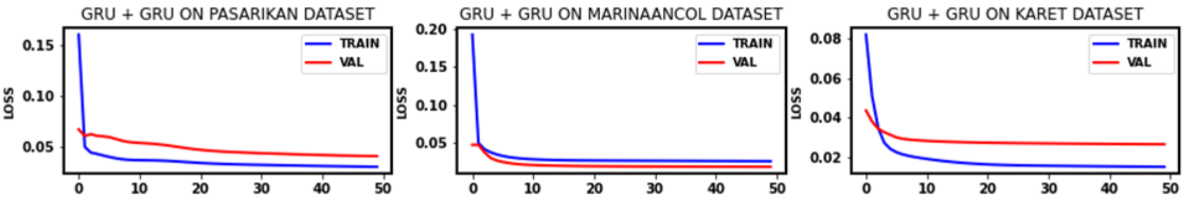

Fig 6. Training and validation losses of stacked GRU on three different datasets

Table 1. Flood Alert Classification

Hyperparameter	Value
Number of Neurons/layer (LSTM/GRU)	32
Number of Neurons (Dense layer)	7
Optimizer	Adam
Loss	Mean Squared Error
Batch Size	32
Number of epochs	50

For the purpose of unnessary repetition here we only report the performance of the stacked GRU model to see how close the predicted water level values against the known water level values. Figure 7, 8, and 9 respectively show side by side the normalized water level predicted by the model versus the corresponding known water level values for three different datasets. As can be seen from the figures, the predicted water level is able to properly track the 329 known water level values. In line with Figure 6 the water level values predicted by the model on Marina Ancol dataset, shown by Figure 8, yield the lowest prediction error.

Table 2 shows the summary of the performance of the three models on three different datasets. The water level prediction errors generated by each model is calculated using three different evaluation schemes namely root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). In spite of the three models show somewhat similar performance on three different datasets, apparently, from Table 3 it can be seen that the stacked GRU model has a slightly better overall prediction errors on all datasets.

Table 2. Water level **p**rediction errors of the models on 3 different datasets (in mm)

Floodgates	Model	RMSE	MAE	MAPE
Karet	LSTM	697.23	345.46	0.10
	GRU	694.51	326.56	0.09
	LSTM-GRU	698.31	329.87	0.09
Marina Ancol	LSTM	111.47	82.99	0.05
	GRU	109.73	82.91	0.05
	LSTM-GRU	113.64	87.25	0.05
Pasar Ikan	LSTM	186.19	123.58	0.11
	GRU	183.18	120.21	0.11
	LSTM-GRU	190.06	130.60	0.11

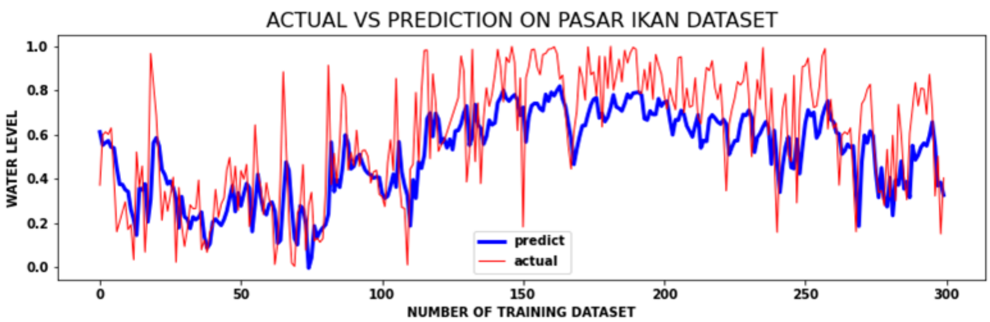

Fig 7. Normalized water level predicted by stacked GRU versus the known water level values on Pasar Ikan dataset

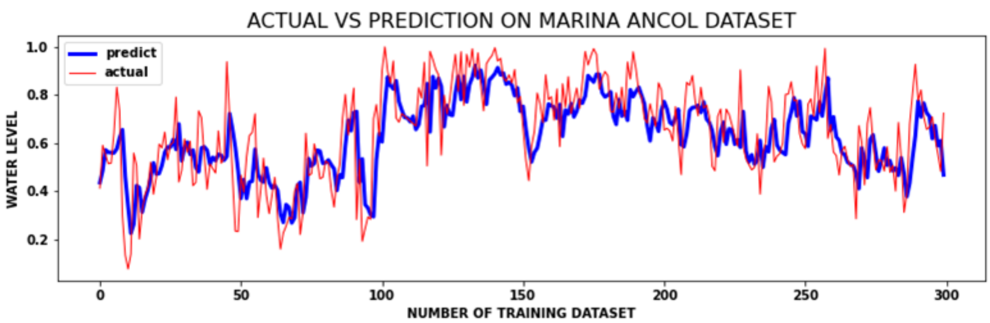

Fig 8. Normalized water level predicted by stacked GRU versus the known water level values on Marina Ancol dataset

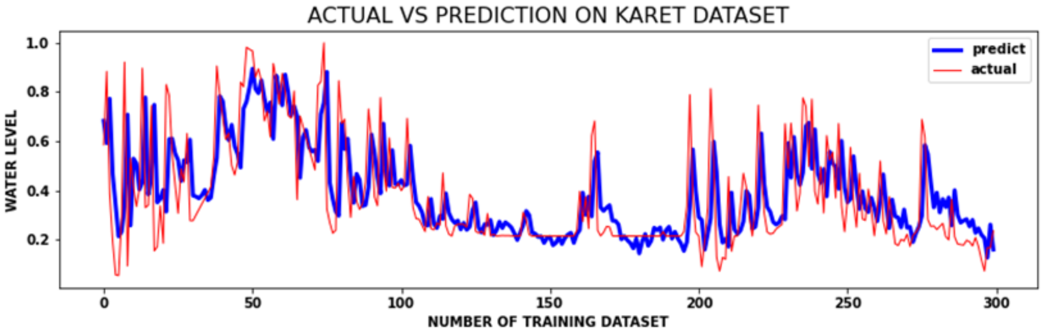

Fig 9. Normalized water level predicted by stacked GRU versus the known water level values on Karet dataset

Table 4. Result of Flood Alert Classification

Water Level	Floodgates	Day-1	Day-2	Day-3	Day-4	Day-5	Day-6	Day-7	Categories
Actual Value	Karet	270.42	264.74	263.36	280.64	264.65	261.56	261.32	Alert 4
Predicted		302.18	305.66	298.74	308.74	307.83	297.71	313.09	
Actual Value	Marina Ancol	169.61	180.62	174.91	169.27	169.27	180.4	185.18	Alert 3
Predicted		173.20	175.84	174.44	176.01	174.43	173.74	175.69	
Actual Value	Pasar Ikan	169.54	183.30	178.97	172.93	169.68	179.12	182.76	Alert 4
Predicted		165.11	167.80	165.81	166.91	164.51	164.72	162.32	

After succeeding to get predicted water level for each crucial floodgate with sufficient accuracy, we then utilize the predicted water level to classify the flood alert in each corresponding floodgate. Table 3 shows the flood alert categories for three different floodgates. As it is observed, each floodgate has its own base criteria to categorize the flood alert depending on the depth of the water basin. Based on Table 3, eventually we will able to make flood prediction on the basis of flood alert categories shown in Table 4.

Table 3. Flood Alert Classification (in cm)

Categories	Karet	Pasar Ikan	Marina Ancol
Alert 1	> 600	> 250	> 250
Alert 2	550-600	200-250	200-250
Alert 3	450-550	170-200	170-200
Alert 4	< 450	< 170	< 170

5. Conclusion and Future Work

An attempt to develop a base model of Jakarta flood early warning services (FEWS) using RNN algorithms has been successfully carried out. Using three different concatenated datasets, each of RNN based models namely stacked GRU-GRU, stacked LSTM-LSTM, and stacked LSTM-GRU succeeds to predict water level in 7 days ahead with a considerable good accuracy on three crucial city floodgates. The recursive method implemented in "walk-forward validation" strategy is successfully employed to perform multi-step prediction. In spite of all the models are capable of giving low prediction errors, however, the stacked GRU-GRU model is somewhat better than the other two models. The stacked GRU with 32 neuron each is capable of yielding RMSE, MAE, and MAPE of 109.73, 82.91, and 0.05 respectively on Marina Ancol dataset.

Computational experiments carried out in this study recognize many imperfections and may be improved further to get more accurate results. Collecting much more historical data on local weather and floodgate water level may significantly improve the accuracy of the prediction. Moreover, utilizing more complex architecture such as encoder-decoder LSTM may be able to considerably reduce the prediction errors.

Author contributions

Hanis A. Saputri: Data curation, Training, validation, and testing the models, Writing-Original draft preparation.
Diaz D. Santika: Data curation, Conceptualization, Methodology, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1]. M. Garschagen, G. A. K. Surtiari, and M. Harb, "Is Jakarta's new flood risk reduction strategy transformational?," *Sustain.*, vol. 10, no. 8, pp. 4–7, 2018, doi: 10.3390/su10082934.

[2]. Gill, D. R. . (2022). A Study of Framework of Behavioural Driven Development: Methodologies, Advantages, and Challenges. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(2), 09–12. https://doi.org/10.17762/ijfrcsce.v8i2.2068

[3]. M. Arsyad, "Analisis parameter curah hujan trhdp fenomena perubahan iklim," pp. 139–145, 2021.

[4]. A. Faruq, S. S. Abdullah, A. Marto, M. A. A. Bakar, Samin, and A. Mubin, "River water level forecasting for flood warning system using deep learning long short-term memory network," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 821, no. 1, 2020, doi: 10.1088/1757-899X/821/1/012026.

[5]. Kabisha, M. S., Rahim, K. A., Khaliluzzaman, M., & Khan, S. I. (2022). Face and Hand Gesture Recognition Based Person Identification System using Convolutional Neural Network. International Journal of Intelligent Systems and Applications in Engineering, 10(1), 105–115. https://doi.org/10.18201/ijisae.2022.273

[6]. M. Chhetri, S. Kumar, P. P. Roy, and B. G. Kim, "Deep BLSTM-GRU model for monthly rainfall prediction: A case

study of Simtokha, Bhutan," *Remote Sens.*, vol. 12, no. 19, pp. 1–13, 2020, doi: 10.3390/rs12193174.

[7]. Krishna, P. R. ., and P. . Rajarajeswari. "EapGAFS: Microarray Dataset for Ensemble Classification for Diseases Prediction". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 8, Aug. 2022, pp. 01-15, doi:10.17762/ijritcc.v10i8.5664.

[8]. M. Cho, C. Kim, K. Jung, and H. Jung, "Water Level Prediction Model Applying a Long Short-Term (LSTM)-Gated Recurrent Unit (GRU) Method for Flood Prediction," *MDPI*, vol. 14, no. 2221, 2022.

[9]. D. E. Tarkus, S. R. U. A. Sompie, and A. Jacobus, "Implementasi Metode Recurrent Neural Network pada Pengklasifikasian Kualitas Telur Puyuh," *J. Tek. Inform.*, vol. 15, no. 2, pp. 137–144, 2020.

[10]. I. Cholissodin and A. A. Soebroto, *AI , MACHINE LEARNING & DEEP LEARNING (Teori & Implementasi)*. Malang, 2020.