Forecasting the number of dengue fever based on weather conditions using ensemble forecasting method

Mursyidatun Nabilah, Raras Tyasnurita, Faizal Mahananto, Wiwik Anggraeni, Retno Aulia Vinarti, Ahmad Muklason

Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Article Info

ABSTRACT

Dengue fever is still a crucial public health problem in Indonesia, with the highest case fatality rate (CFR) is 1.01% in East Java, Malang Regency. One of the solutions to control the death rate and cases is to forecast the cases number. This study proposed ensemble forecasting that build from several penalized regressions. Penalized regressions are able to overcome linear regression analysis' shortcomings by using penalty values, that will affect regression's coefficient, resulting on regression model with a slight bias in order to reduce parameter estimations and prediction values' variances. Penalized regressions are evaluated and built as ensemble forecasting method to minimize the shortcomings of other existing model, so it could produce more accurate values comparing to single penalized regression model. The result showed that the ensemble model `consists of smoothly clipped absolute deviation (SCAD) and Elastic-Net is sufficient to capture data patterns with root mean squared error (RMSE) 6.38.

Corresponding Author:

Wiwik Anggraeni
Department of Information Systems, Institut Teknologi Sepuluh Nopember
Sukolilo, Surabaya, Jawa Timur 60111, Indonesia
Email: wiwik@is.its.ac.id

1. INTRODUCTION

Dengue fever is a dangerous infectious disease whose cases have steadily increased over years. Dengue fever is caused by a virus that is in the saliva of Aedes mosquito that injects human body parts that varies from mild into severe conditions [1], [2]. As stated from Epidemiological data and Surveillance Center, Ministry of Health, Indonesia, in Indonesia, dengue fever is still a crucial problem, this is because the number of infections and the area of distribution is increasing along with the increase in mobility and population density.

Based on the Ministry of Health Republic of Indonesia's data, in 2019, the case fatality rate (CFR) of dengue fever showed a value of 0.67% on a national scale. CFR is obtained from the proportion of deaths to all reported cases. A province is said to have a high CFR if it exceeds 1%. One of the provinces that has a high CFR is East Java with 1.01%. Based on the Malang Regency Health Office, Malang Regency is the area with the highest number of cases and deaths from dengue fever in East Java in 2019, therefore efforts are needed to control the death rate from dengue fever in Malang.

To control the mortality rate in Malang Regency, one of the efforts is to predict the number of dengue fever cases in the future, one of the research conducted is building a model to forecast so that the parties in charge could take steps and arrange policies to minimize the increase in cases and mortality rates. Several forecasts related to dengue fever have been carried out by utilizing weekly or monthly number of cases [3]–[5]. Based on the previous research, there is a fairly high correlation between the number of cases of dengue fever with rainfall, temperature [6], and humidity [7].

Penalized regression is a regression model using penalty that aims to reduce overfitting in multiple linear regression [8]. In this study, Ridge, Lasso, Elastic Net, smoothly clipped absolute deviation (SCAD), and minimax concave penalty (MCP) were explored. In order to overcome limitations of single forecasting model, ensemble methods are able to increase the performance of base model with higher accuracy and identify complex object, and uncertainties [9]–[12].

2. METHODOLOGY

Based on Figure 1, there are four big steps to build ensemble model with penalized regression. First, raw data are gathered from various sources. It consists of climate data (temperature, humidity, wind speed, rainfall) and number of dengue cases. Data cleaning is carried out to produce processed data that are ready to be used for model development.
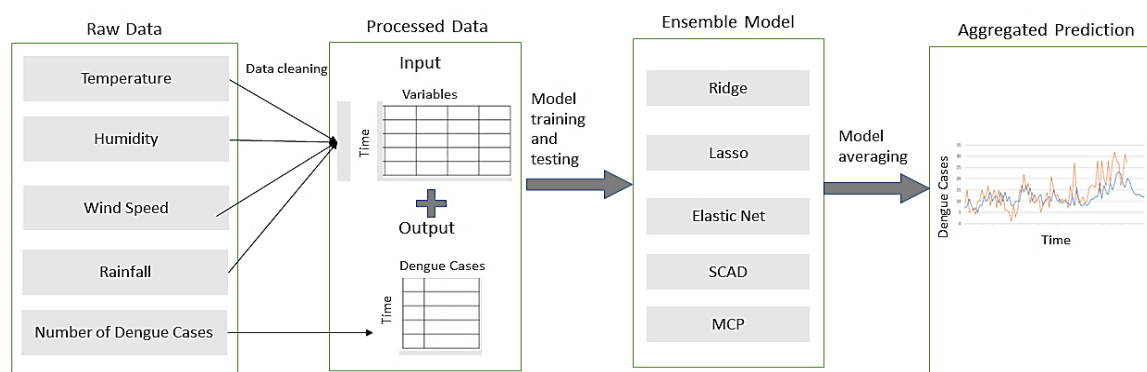

Figure 1. Model architecture

After raw data collection and cleaning stage, the next steps are test the correlation between variables. Data will be splitted into two parts that consist of training data and testing data. Training data will be used to train the model and the testing one will be used to measure model performances, also those datas will be used to determine parameter of penalized.

In building ensemble forecasting model, five penalized regressions that consist of Ridge, Lasso, Elastic Net, SCAD, and MCP will be trained and validated by each. Ridge regression widely used for high dimensional data where independendent variables are highly correlated, this method aims to reduce multicollinearity [13], Lasso is a method that used regularization and variable selection to increase interpretability and accuracy [14], Elastic Net is a combination of Ridge and Lasso regressions, so it will retain the advantage of both methods [15], SCAD regression aims to improve Lasso's penalty by reducing the bias in the model because the Lasso penalty tends to be linear in the size of the regression coefficient [16], and MCP is other alternative to give less biased variables in sparse model [17]. Penalized regression parameter will be determined.

After evaluating each model, aggregated prediction is formed by calculating the average prediction results from the model (averaging). In general, the steps carried out are implemented by taking sequential data based on the time dimension. After the ensemble forecasting equation has been successfully formed, then forecasting is carried out on the dependent variable (weekly number of cases of dengue fever) using the ensemble forecasting model that has been formed on the test data. After the formation of forecasting models and predictions have been made, the analysis is carried out by predicting the magnitude of the incidence of dengue fever and strategy analysis. After that, the model performance test was carried out.

This study will try to test two forms of data to get the most optimal model results. The form of data to be tested consists of normal data and data that has been transformed into natural logarithm (ln). Based on the research, the natural logarithm transformation was carried out to stabilize the variance when performing standard regression procedures. In addition to the unstable variance (not constant), the transformation can also be used to correct for non-linearity and residuals that are not normally distributed (non-normality) [18].

3. RESULTS AND DISCUSSION

The experiments were performed on an Intel® Core™ i5-7200U central processing unit (CPU) @ 2.50 GHz 2.70 GHz, random-access memory (RAM) with 8 GB (gigabyte) which is running on Windows 10 home

single language x64 bit. The software tool used is Rstudio and R for the programming language. The steps show the results from each research steps that consist of splitting data, determining parameters, building penalized regression model, ensembling model, and compare the model's performance with other related methods.

3.1. Splitting data training and data testing

To conduct training on the model, the data is divided into two parts into training data and testing data. Training data is part of a dataset that is trained to make predictions or perform functions from other machine learning algorithms according to their respective goals. Basically, the user provides clues through an algorithm so that the trained machine can find the correlation on its own. While data testing is part of the dataset that is tested to see the accuracy of the model, in other words, its performance. The distribution of overall data from 2014 to 2018 sequentially with the proportion of training data compared to testing data of 70% and 30%.

3.2. Determined penalized regression parameter

The parameter used in penalized regression is the lambda value. The lambda parameter controls the amount of regularization applied to the regression model. The larger the lambda value, the more coefficients are depreciated to zero. When the lambda value is equal to 0, the regularization does not apply and the model runs linear regression. Lambda value with cross validation score with the smallest error value are taken for each model and the proportion of data [19]. Table 1 shows the selected lambda values along with the lowest mean squared error (MSE) score for each lambda in the Ridge, Lasso, and Elastic Net models. When compared with the selected lambda, Lasso and Elastic Net have a large enough lambda value from the Ridge model. This is because if the lambda value is greater, there is a possibility that a variable has a coefficient equal to zero. It can be said that several independent variables are not chosen to be predictors in Lasso and Elastic Net regression models, considering that Lasso regression has a variable selection feature in it [20], and Elastic Net is a combination of Ridge and Lasso models [21].

The calculation of the best lambda values for the SCAD and MCP models in Table 2 is slightly different from Ridge, Lasso, and Elastic Net models. The best lambda value is selected based on the lowest cross-validation error (CVE) value. The best lambda that can be used on SCAD is 0.908 with a CVE of 73.17, and MCP has a lambda of 0.520 with a CVE of 69.34.

Table 1. Ridge, Lasso, and Elastic Net's lambda values

Data proportion	Method	Lambda values	MSE
70:30	Ridge	1.696	123.49
70:30	Lasso	5.060	121.16
70:30	Elastic net	6.975	117.42

Table 2. SCAD and MCP's lambda values

Data proportion	Method	Lambda values	CVE
70:30	SCAD	0.908	73.17
70:30	MCP	0.520	69.34

3.3. Building penalized regression model

Testing data is carried out in each penalized regression model with the proportion of 70:30 for training and testing data. The performance of each model is measured based on the root mean squared error (RMSE) and symmetric mean absolute percentage error (SMAPE) values. The model was tested on both forms of data, namely normal and logarithmic transformation data. Since the number of cases' smallest error numbers on normal data (RMSE: 6.38) is lower than logarithmic transformation (RMSE: 8.95), normal data is chosen for building the penalized regression model. The performance results of each penalized regression on normal data can be seen in Table 3.

Table 3. Penalized regression's performances

Method	RMSE	SMAPE
Ridge	6.67	42%
Lasso	7.04	39%
Elastic Net	6.51	41%
SCAD	6.45	37%
MCP	6.70	39%

Based on the test results on normal data, the SCAD model has the best performance among other penalized regression models, followed by the Elastic Net, Ridge, MCP, and finally Lasso models. Based on the order of the smallest RMSE, the models will be combined based on scenarios based on the best RMSE value. When it is viewed from the prediction pattern of each model, Ridge on Figure 2 can capture the pattern quite well, it can be seen from the prediction pattern which tends to follow the increase and decrease in the actual data.

The SCAD model in Figure 3 can also capture data patterns well. When compared to the Ridge and MCP models, SCAD tends to be more able to follow patterns in the early period with time range of October 28, 2017 (10/28/2017) to December 28, 2017 (12/28/2017). It can be seen from the pattern of predictive data that tends to decrease so that the range of error values is smaller in this section. Even so, the increase in data that occurred in the period from October 31, 2018 (10/31/2018) to November 30, 2018 (11/30/2018) could not follow the pattern as good as the Ridge and MCP models.
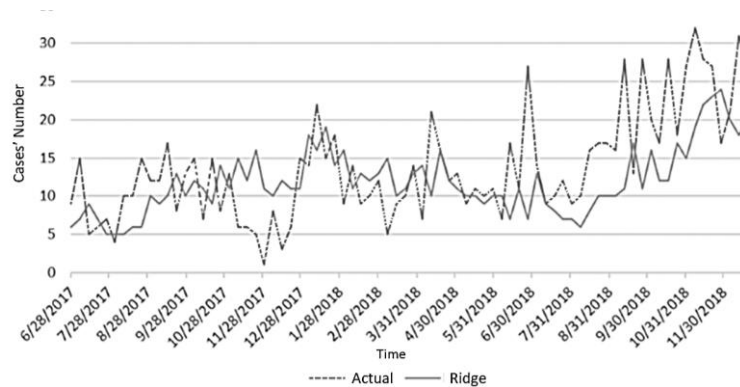
Figure 2. Comparison between ridge vs actual
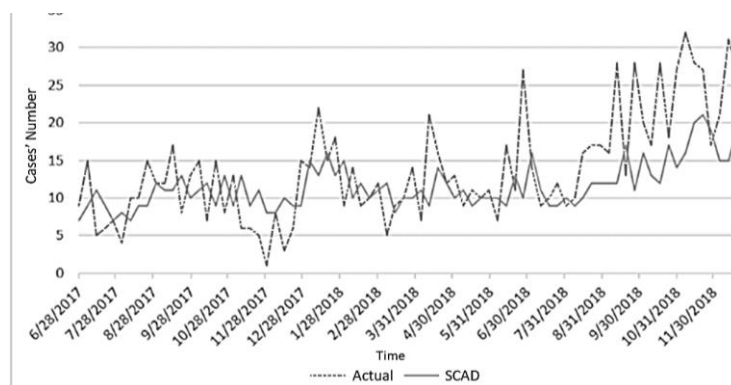
Figure 3. Comparison between scad vs actual

The MCP model in Figure 4 can also follow the data pattern quite well. It can be seen from the ability of the prediction results to follow the actual value. Forecasting using Lasso has a variable selection feature, where the model will select independent variables that have relevance to the dependent variable. Even so, the results of the Lasso model are less able to capture the actual data pattern and tend to be less sensitive. It can be seen in Figure 5 where the increase in cases cannot be captured properly by the Lasso model, so this can also be the cause of the RMSE of the Lasso model having the greatest value among other models. The prediction of the Elastic Net model in Figure 6 also has a variable selection feature like Lasso's. However, Elastic Net can still capture patterns and spikes in the testing data well, this is because the model also combines the Ridge model in it, so that a combination of Ridge predictions is obtained that can handle multicollinearity [22] and could select variables according to existing data patterns.

3.4. Building ensemble model

Based on the best RMSE value of the model that has been tested further, a combination of each is carried out according to the scenario that has been set in. The normal data ensemble scenario in Table 4 shows

that the BEST II scenario with the combination of SCAD + Elastic Net models has the lowest RMSE, that means the model has the lowest error compared to others. And then based on low to high RMSE, the performance of the model followed by BEST III, ALL, and lastly BEST IV scenarios.
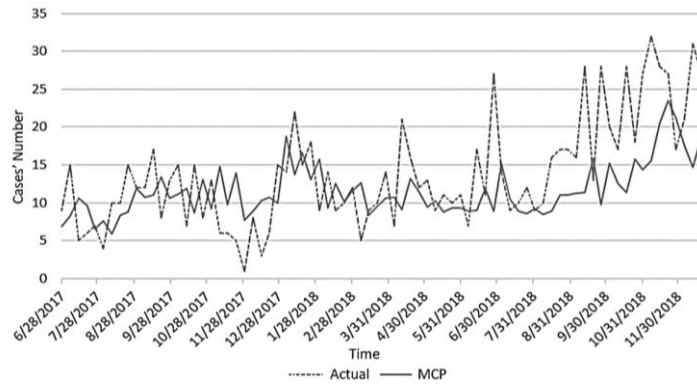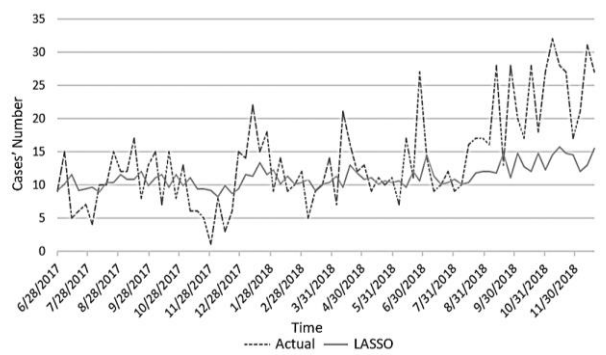

Figure 4. Comparison between MCP vs actual

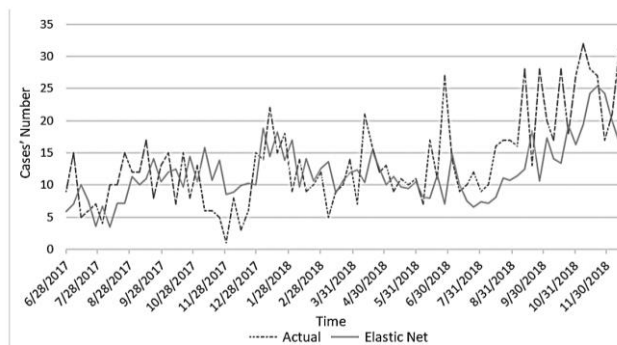
Figure 5. Comparison between Lasso vs actual

Figure 6. Comparison between elastic net and actual

Table 4. Performances of ensemble model scenario

Scenario	Model combination	RMSE	SMAPE
BEST II	SCAD + Elastic Net	6.38	39%
BEST III	SCAD + Elastic Net + Ridge	6.45	39%
BEST IV	SCAD + Elastic Net + Ridge + MCP	6.51	39%
ALL	SCAD + Elastic Net + Ridge + MCP + LASSO	6.50	38%

The best model from the experimental results is used to predict the next 8 weeks, starting from January 2019 to February 2019. To predict the number of dengue fever cases in the next 8 weeks, the main thing to do is to predict each independent variable first, such as temperature, humidity, rainfall and wind speed. In the independent variable forecasting process, the methods used are different depending on the data pattern. Based on observations, the variables of air temperature, air humidity, and wind speed have cyclical data patterns, where the data patterns are repeated over a long period of time [23]. Therefore, the three variables can be predicted using the multiplicative decomposition method [24]. The results of temperature forecasting for the next 8 weeks are shown in Figure 7.
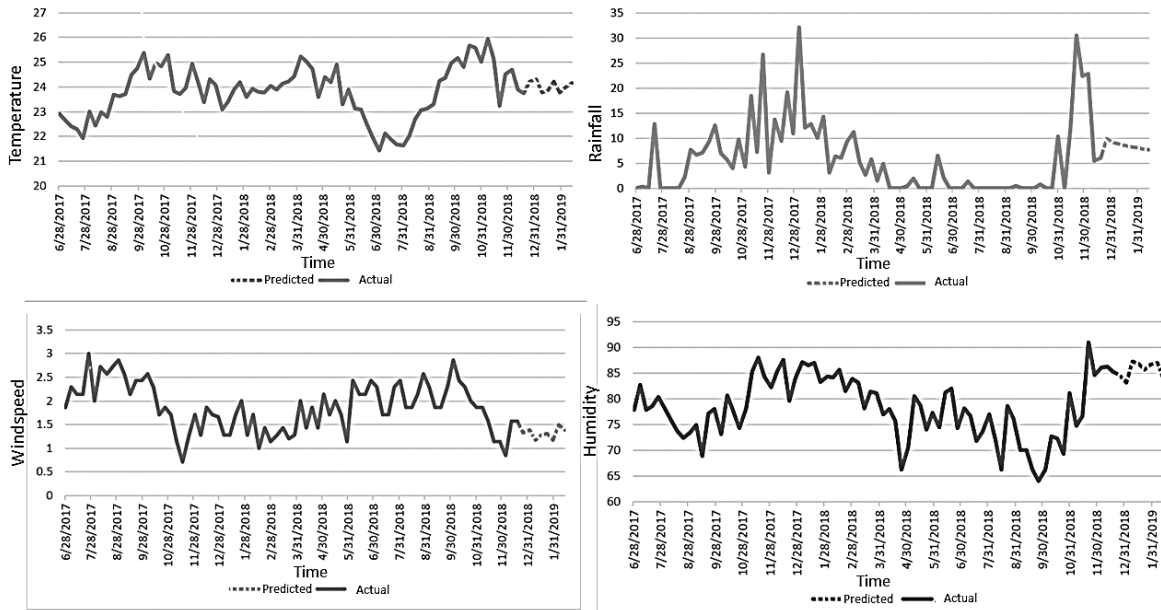
Figure 7. Forecast of climate data

The forecast used the best model by predicting the number of cases each week ahead, and the previous forecasting results are used as a lag feature in the next data. This is done 8 times until the forecasting results are formed. The combination of SCAD + Elastic Net with normal data which is the model with the best performance is used to predict the number of cases of dengue fever in Malang Regency. Forecasting period is the first 8 weeks of the year, from January 2019 to February 2019. From the forecasting results obtained, there will be a decrease in the number of dengue fever cases for the next 8 weeks. This visualization is shown in Figure 8.
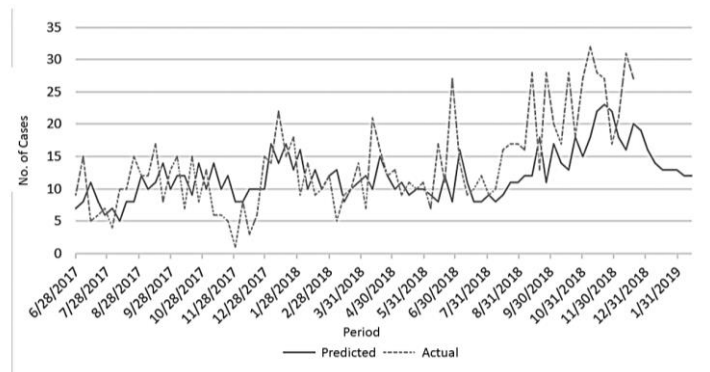

Figure 8. Forecast of dengue cases' numbers in 8 weeks ahead

When compared with the actual data on dengue fever cases in 2019 listed in Table 5, the number of cases in 2019 tends to increase. The cause of these differences can be caused by the presence of other factors or variables that cause an increase in cases of dengue fever. The dengue transmission of dengue cases in East Java tends to be influenced by population density, population mobility, urbanization, residental areas and in public places [25]. In this research, the variable used to predict is only based on climate, so it is possible that there are other factors than climate that are sufficient to have more influence on the increase or decrease in the number of dengue cases.

Table 5. Predicted vs actual dengue cases of 2019

Week number	Actual cases of 2019	Predicted cases of 2019
1	16	19
2	23	16
3	26	14
4	27	13
5	28	13
6	35	13
7	42	12
8	27	12

3.5. Comparison with other methods

To find out more about whether the forecasting method using the SCAD + Elastic Net method that has been carried out is good enough, the forecasting results need to be compared with other methods. Comparison is generated by the SCAD + Elastic Net model with another method, namely multiple linear regression [26]. The models were compared based on the RMSE value where performances are contained in Table 6.

Table 6. Ensemble model vs multiple linear regression

Model	RMSE	SMAPE
SCAD + Elastic Net	6.38	39%
Multiple linear regression	6..45	41%

In addition, in terms of determining the regression coefficient of the independent variable, the Elastic Net and SCAD models which are part of penalized regression have the ability to reduce the regression coefficient value to 0. In other words, it can eliminate independent variables that are less significant to the model [27]. In multiple linear regression, all independent variables such as wind velocity, rainfall, humidity, air temperature, lag-1, and intercept (the mean value of the response variable when all predictor equals to zero) are considered in the model development. It is different with the Elastic Net model, which is only 1 variable was selected, namely lag-1, lag-1 consists of the number of dengue cases that were pushed back one day from the original data. While in the SCAD model, humidity was eliminated from the model. These results can be seen in Table 7.

Table 7. regression coefficient between multiple linear regression, SCAD, and Elastic Net

Model	Wind velocity	Rainfall	Humidity	Temperature	Lag-1	Intercept
Multiple linear regression	-1.9127	0.21351	0.00864	0.913607	0.51028	-14.893
Elastic Net	0	0	0	0	0.23306	8.0364
SCAD	-1.9559	0.21615	0	0.9094702	0.51088	-14.047

4. CONCLUSION

Based on the results of the research that has been done, the following conclusions can be drawn: The best method for forecasting dengue fever cases is the ensemble model using a combination of SCAD + Elastic Net finalized regression with RMSE of 6.38. The logarithm transformation of the data on the number of cases does not provide better performance than normal data. It can be seen from the smallest RMSE value of the data from the ln transformation is 8.95 and for the normal data is 6.38. Based on the results of variable selection from one of ensemble forming models (Elastic Net), only the lag-1 variable has a regression coefficient that is not equal to 0, it means that in the Elastic Net regression, only the lag-1 variable is used in constructing model. While in the SCAD, there is only one variable has a regression coefficient that is equal to 0. In order to improve the forecast performance, the selection of variables need to be reconsidered. In addition to the climate factors

such as temperature, humidity, rainfall and wind speed, other variables can be explored for future research such as population density, population mobility, economic growth, environmental sanitation, urbanization, and community behavior.

REFERENCES

[1] N. Khetarpal and I. Khanna, "Dengue fever: causes, complications, and vaccine strategies," *Journal of Immunology Research*, vol. 2016, 2016, doi: 10.1155/2016/6803098.

[2] S. A. alias Balamurugan, M. S. M. Mallick, and G. Chinthana, "Improved prediction of dengue outbreak using combinatorial feature selector and classifier based on entropy weighted score based optimal ranking," *Informatics in Medicine Unlocked*, vol. 20, 2020, doi: 10.1016/j.imu.2020.100400.

[3] N. A. Lestari, R. Tyasnurita, R. A. Vinarti, and W. Anggraeni, "Long short-term memory forecasting model for dengue fever cases in Malang Regency, Indonesia," *Procedia Computer Science*, vol. 197, pp. 180–188, 2021, doi: 10.1016/j.procs.2021.12.131.

[4] A. L. Buczak, B. Baugher, L. J. Moniz, T. Bagley, S. M. Babin, and E. Guven, "Ensemble method for dengue prediction," *PLoS ONE*, vol. 13, no. 1, 2018, doi: 10.1371/journal.pone.0189988.

[5] C. M. Benedum, K. M. Shea, H. E. Jenkins, L. Y. Kim, and N. Markuzon, "Weekly dengue forecasts in Iquitos, Peru; San Juan, Puerto Rico; and Singapore," *PLoS Neglected Tropical Diseases*, vol. 14, no. 10, pp. 1–26, 2020, doi: 10.1371/journal.pntd.0008710.

[6] Y. Li, Q. Dou, Y. Lu, H. Xiang, X. Yu, and S. Liu, "Effects of ambient temperature and precipitation on the risk of dengue fever: a systematic review and updated meta-analysis," *Environmental Research*, vol. 191, 2020, doi: 10.1016/j.envres.2020.110043.

[7] A. Susilawaty, R. Ekasari, L. Widiastuty, D. R. Wijaya, Z. Arranury, and S. Basri, "Climate factors and dengue fever occurrence in Makassar during period of 2011–2017," *Gaceta Sanitaria*, vol. 35, pp. S408–S412, 2021, doi: 10.1016/j.gaceta.2021.10.063.

[8] H. Kumar and B. K. Hooda, "Comparison of penalized and multiple linear regression for prediction of milk yield in crossbred cattle," *International Journal of Agricultural and Statistical Sciences*, vol. 11, no. 1, pp. 151–154, 2015.

[9] R. Walambe, A. Marathe, K. Kotecha, and G. Ghinea, "Lightweight object detection ensemble framework for autonomous vehicles in challenging weather conditions," *Computational Intelligence and Neuroscience*, vol. 2021, 2021, doi: 10.1155/2021/5278820.

[10] R. Walambe, A. Marathe, and K. Kotecha, "Multiscale object detection from drone imagery using ensemble transfer learning," *Drones*, vol. 5, no. 3, 2021, doi: 10.3390/drones5030066.

[11] P. Guo *et al.*, "An ensemble forecast model of dengue in Guangzhou, China using climate and social media surveillance data," *Science of the Total Environment*, vol. 647, pp. 752–762, 2019, doi: 10.1016/j.scitotenv.2018.08.044.

[12] Y. Zhu, "Ensemble forecast: a new approach to uncertainty and predictability," *Advances in Atmospheric Sciences*, vol. 22, no. 6, pp. 781–788, 2005, doi: 10.1007/BF02918678.

[13] S. Liu and E. Dobriban, "Ridge regression: structure, cross-validation, and sketching," 2019.

[14] F. Emmert-Streib and M. Dehmer, "High-dimensional Lasso-based computational regression models: regularization, shrinkage, and selection," *Machine Learning and Knowledge Extraction*, vol. 1, no. 1, pp. 359–383, 2019, doi: 10.3390/make1010021.

[15] Y. Liu, Z. Yu, C. Chen, Y. Han, and B. Yu, "Prediction of protein crotonylation sites through LightGBM classifier based on SMOTE and Elastic Net," *Analytical Biochemistry*, vol. 609, 2020, doi: 10.1016/j.ab.2020.113903.

[16] G. Aneiros, S. Novo, and P. Vieu, "Variable selection in functional regression models: a review," *Journal of Multivariate Analysis*, vol. 188, 2022, doi: 10.1016/j.jmva.2021.104871.

[17] J. You, Y. Jiao, X. Lu, and T. Zeng, "A nonconvex model with minimax concave penalty for image restoration," *Journal of Scientific Computing*, vol. 78, no. 2, pp. 1063–1086, 2019, doi: 10.1007/s10915-018-0801-z.

[18] J. M. Brunkard, E. Cifuentes, and S. J. Rothenberg, "Assessing the roles of temperature, precipitation, and ENSO in dengue re-emergence on the Texas-Mexico border region," *Salud Publica de Mexico*, vol. 50, no. 3, pp. 227–234, 2008, doi: 10.1590/S0036-36342008000300006.

[19] L. E. Melkumova and S. Y. Shatskikh, "Comparing Ridge and Lasso estimators for data analysis," *Procedia Engineering*, vol. 201, pp. 746–755, 2017, doi: 10.1016/j.proeng.2017.09.615.

[20] R. Tibshirani, "Regression shrinkage and selection via the Lasso: a retrospective," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 73, no. 3, pp. 273–282, 2011, doi: 10.1111/j.1467-9868.2011.00771.x.

[21] H. Zou and T. Hastie, "Regularization and variable selection via the Elastic Net," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, Apr. 2005, doi: 10.1111/j.1467-9868.2005.00503.x.

[22] F. Sami, M. Amin, and M. M. Butt, "On the ridge estimation of the conway-maxwell Poisson regression model with multicollinearity: methods and applications," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 1, Jan. 2022, doi: 10.1002/cpe.6477.

[23] P. Verboon and R. Leontjevas, "Analyzing cyclic patterns in psychological data: a tutorial," *The Quantitative Methods for Psychology*, vol. 14, no. 4, pp. 218–234, 2018, doi: 10.20982/tqmp.14.4.p218.

[24] R. J. H. and G. Athanasopoulos, *Forecasting: principles and practice*. 2018.

[25] D. P. Dayani, "The overview of dengue hemorrhagic fever in East Java during 2015-2017," *Jurnal Berkala Epidemiologi*, vol. 8, no. 1, p. 35, Jan. 2020, doi: 10.20473/jbe.v8i12020.35-41.

[26] W. Anggraeni *et al.*, "Modified regression approach for predicting number of dengue fever incidents in Malang Indonesia," *Procedia Computer Science*, vol. 124, pp. 142–150, 2017, doi: 10.1016/j.procs.2017.12.140.

[27] Y.-J. Yoon and M.-S. Song, "Variable selection via penalized regression," *Communications for Statistical Applications and Methods*, vol. 12, no. 3, pp. 615–624, 2005, doi: 10.5351/ckss.2005.12.3.615.

BIOGRAPHIES OF AUTHORS

Mursyidatun Nabilah received her bachelor's degree from Institut Teknologi Sepuluh Nopember majoring on Information Systems in 2021. She was a part of member on Data Engineering and Business Intelligence research group in Information Systems Department. During her time as a student, she was active in organizational activities, committees, and extracurricular such as Himpunan Mahasiswa Sistem Informasi (HMSI) ITS, Information Systems' Expo, and Marching Band. She has research interest on data analysis, machine learning, and forecasting. She can be contacted at email: yusi.nabilah@gmail.com.

Raras Tyasnurita main research interests are artificial intelligence, combinatorial optimisation, machine learning, and forecasting. She is currently a lecturer within the Data Engineering and Business Intelligence research group in Information Systems Department, Institut Teknologi Sepuluh Nopember. She was an Editor in Chief of SISFO Journal (2009-2010), SESINDO 2013 proceeding as a national conference in Information Systems, and International Seminar on Science and Technology (ISST) 2019. She can be contacted at email: raras@is.its.ac.id.

Faizal Mahananto current research interests include healthcare information systems, business intelligence in medicine and healthcare, heart rate variability, and biomedical signal processing. He received the S.Kom. degree from Institut Teknologi Sepuluh Nopember, Indonesia, in 2008 and the M.Eng. and Ph.D. degree from the Kumamoto University, Japan, in 2012 and 2015, respectively. He is currently working in the Department of Information Systems, Institut Teknologi Sepuluh Nopember, Indonesia. He can be contacted at email: fmahananto@gmail.com.

Wiwik Anggraini current research interests are business analytics: modelling, forecating, optimization, and predictive analysis. She received her Bachelor and Master degree from Institut Teknologi Sepuluh Nopember, majoring on Mathematics (1997) and Informatics (2003), respectively. She is a lecturer in Data Engineering and Business Intelligence research group in Information Systems Department. Currently, she is taking a doctorate degree on electrical engineering, Institut Teknologi Sepuluh Nopember. She can be contacted at email: wiwik@is.its.ac.id.

Retno Aulia Vinarti main research interests are Healthcare Informatics, Rule-based Systems, Bayesian Networks, Knowledge-driven Models, and Prediction. She received her bachelor and master's degree in computing from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 2009 and 2011, respectively. She received her PhD degree from Trinity College Dublin, Dublin, Ireland, in 2019. She is currently a Head of Data Engineering and Business Intelligence Laboratory in Information Systems Department, Institut Teknologi Sepuluh Nopember. She can be contacted at email: Zahra_17@is.its.ac.id.

Ahmad Muklason main research interests are Hyper/Meta-heuristics, Examination Timetabling Problems, Combinatorial Optimisations Problems, and Fairness Urban Computing. He received his bachelor's degree in computing from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 2006, Master's degree in Computer and Information Sciences Department from Universiti Teknologi Petronas, Malaysia, in 2009, and PhD degree in Information Technology, School of Computer Science, from The University of Nottingham, United Kingdom, in 2017. He can be contacted at email: mukhlason@is.its.ac.id.