

Comparative Analysis of Ridge, LASSO, and Elastic Net Regularization Approaches in Handling Multicollinearity for Infant Mortality Data in South Sulawesi

Arief Rahman Nur^{1*}, Andi Kresna Jaya^{2*}, and Siswanto Siswanto^{3*}

** Department of Statistics, Hasanuddin University, Indonesia*

Email: ¹ariefrahmannur18@gmail.com, ²andikresna.jaya@gmail.com, ³siswanto@unhas.ac.id

Abstract

Infant mortality rate is a crucial indicator for assessing the health and infant care quality in a region. In the effort to reduce infant mortality rates, regression analysis serves as a tool to identify influential factors. However, regression analysis often encounters the challenge of multicollinearity, which involves high correlation among predictor variables. To address this issue, various regularization techniques can be applied, such as ridge regression, least absolute shrinkage and selection operator (LASSO), and elastic net. Ridge regression aims to control coefficient variance, while LASSO directs some coefficients to zero, functioning as variable selection. Elastic net combines the strengths of both methods by merging ridge and LASSO regularization. The objective of this research is to evaluate the performance of ridge regression, elastic net, and LASSO methods in handling multicollinearity issues, utilizing infant mortality rate data in South Sulawesi Province. The results indicate that the elastic net method outperforms both Ridge and LASSO methods. The best-performing model is obtained through elastic net with a coefficient of determination value of 60.81%, whereas ridge and LASSO methods yield coefficient of determination values of 54.11% and 58.18%, respectively. This demonstrates that the application of the elastic net method is capable of producing more accurate results in modeling the variables within the analysis of infant mortality rate data compared to other methods.

Keywords: elastic net, infant mortality, LASSO, multicollinearity, regularization, ridge

1. INTRODUCTION

The Infant Mortality Rate (IMR) measures the number of infants who die before the age of one year per 1000 live births in one year [7]. This figure reflects public health and is a target of the Sustainable Development Goals (SDGs) [11]. A low infant mortality rate indicates good community health and is associated with child survival as well as social, economic, and environmental factors [14]. Regression analysis is used to identify the factors that influence the infant mortality rate. Multicollinearity occurs when predictor variables in multiple regression analysis have a significant relationship with each other, which can result in the model having high variance [13]. This can interfere with the accuracy of model predictions and result in errors in decision making. To



overcome multicollinearity problems, the method that is often used is the regularization regression method [5].

The research conducted by Altalbany 2021 compared ridge regression, Lasso, and elastic net methods using simulation data in various sample sizes. The results showed the superiority of the elastic net in overcoming multicollinearity at all sample levels, except for data with severe multicollinearity and samples less than 10,000 observations, where Lasso performed better [1]. In addition to simulations, the elastic net was also applied to poverty rate data in Central Java Province, with results showing better performance with a lower Mean Squared Error (MSE) compared to ridge and LASSO, especially in economic and social data modeling [8]. This study initially focused on simulations and socio-economic aspects. However, the plan will be expanded by applying the elastic net method to health data, specifically in the context of Infant Mortality Rate (IMR), to understand the impact of predictor variables on the IMR.

The high infant mortality rate in Indonesia requires more attention to several key factors such as exclusive breastfeeding, infant health services, provision of vitamin A, and the condition of the mother during pregnancy and childbirth. However, these factors are interrelated and potentially cause multicollinearity in the data [11]. To address this, regularization regression methods are techniques used to prevent multicollinearity in regression analysis. This technique involves adding a penalty term to the objective function of linear regression. There are two types of regularization commonly used, namely LASSO, ridge, and elastic net [16]. This technique helps to find the variables that have the most influence on the infant mortality rate by eliminating insignificant variables and reducing the effects of interrelated variables. This study also involves comparing the performance of three regularization regression methods, namely LASSO, ridge, and elastic net, on infant mortality rate data [2]. By comparing these three methods, the study aims to determine the most effective method in identifying the most significant factors influencing the infant mortality rate in South Sulawesi. Thus, this research not only helps to address the issue of multicollinearity but also provides insights into which method is most suitable for analyzing complex infant mortality rate data with diverse influencing factors.

2. METHODOLOGY

The data used in this study is secondary data obtained from the South Sulawesi Province Health Profile 2020 and 2021 published by the South Sulawesi Provincial Health Office. The response variable and the predictor variables in this study are based on the data of 24 Cities/Districts in South Sulawesi. The response variable is the infant mortality rate (Y_1), while the predictor variables are the number of babies given exclusive breastfeeding (aged 6 months to 1 year) (X_1), the increasing number of infant health services (X_2), the number of babies given vitamin A (X_3), the number of low birth weight babies (X_4), and the number of mothers assisted by medical personnel during childbirth (X_5). The method used in this study is ridge regression, LASSO, and elastic net.

2.1 Regression Analysis

The multiple linear regression model is a development of the simple linear regression model, with more than one response variable and only one dependent variable. Multiple linear regression equations that include two or more response variables have a different general form compared to simple linear regression models that only have one response variable [10]. The general form of the multiple linear regression equation is as follows:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon \quad (2.1)$$

If you use sample data, you can use the estimation model from Equation (2.1) as follows:

$$Y = X\beta + \varepsilon \quad (2.2)$$

Based on Equation (2.2), the matrix form of the multiple linear regression model can be written as follows:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{n1} & X_{2n} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

where: \mathbf{Y} is the response variable vector of size $n \times 1$, \mathbf{X} is the independent variable matrix of size $n \times (p + 1)$, $\boldsymbol{\beta}$ is the vector of parameters to be estimated in size $(p + 1) \times 1$, and $\boldsymbol{\varepsilon}$ is a random error vector of size $n \times 1$.

Parameter estimates can be obtained using the least squares method, namely by minimizing the sum of squared errors:

$$S(\boldsymbol{\beta}) = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \quad (2.3)$$

To obtain a least squares method estimator that minimizes equation (2.3), it is required that:

$$\begin{aligned} \frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} &= 0 \\ -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= 0 \\ \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y}) \end{aligned}$$

2.2 Multicollinearity

Multicollinearity is a condition where there is a perfect or near-perfect linear relationship between independent variables in a regression model. Detecting multicollinearity is done by examining the VIF (Variance Inflation Factor) values using the following equation [15]:

$$VIF = \frac{1}{(1-R_i^2)} \quad (2.4)$$

If the obtained values $VIF \geq 10$ and *tolerance* < 0.1 then there is a strong multicollinearity among the response variables and vice versa.

2.3 Cross Validation

Perform cross validation to find the most optimal values of α and λ . Cross validation is a method used to assess the accuracy of the model by dividing the data into several segments. The cross validation process involves splitting the data into two groups: one for model building and the other for validation. This division of data is done randomly to ensure that each observation has an equal chance of being part of the validation dataset [8]. The formula used is as follows [9]:

$$CVE_{(k)} = \frac{1}{k} \sum_{i=1}^k RMSE_i \quad (2.5)$$

2.4 Ridge regression

Ridge regression is a method developed to mitigate the impact of multicollinearity on regression coefficient values. Ridge regression is a modification of the Ordinary Least Squares (OLS) method, which produces more stable coefficient estimates by reducing the variability of coefficients. Although the estimation of regression coefficients becomes slightly biased, this method can still approximate the true parameter values [3]. The steps for ridge regression analysis are as follows:

- Selecting the optimal λ value through cross-validation.
- Performing ridge regression analysis with the optimal λ value using the formula

$$\hat{\boldsymbol{\beta}}_{ridge} = \arg \min \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p \beta_j^2 \right\} \quad (2.6)$$

2.5 Least Absolute Shrinkage and Selection Operator

Least Absolute Shrinkage and Selection Operator (LASSO) method is a development aimed at variable selection while estimating regression parameters. LASSO works by shrinking the regression coefficients of variables that have high correlation with the error, thereby pushing some of these coefficients towards or even to zero. Thus LASSO serves as a variable selection tool while also addressing multicollinearity issues in regression analysis. The steps for LASSO regression analysis are as follows [12]:

- a. Selecting the optimal λ value through cross-validation.
- b. Performing Lasso regression analysis with the optimal λ value using the formula

$$\hat{\beta}_{lasso} = \arg \min \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (2.7)$$

2.6 Elastic net

The elastic net method is an approach in regression analysis that combines principles from ridge regression and Lasso regression [17]. Elastic net applies both types of penalties, namely L1 and L2, to the loss function, enabling this method to effectively address multicollinearity issues and perform relevant feature selection simultaneously [6]. With this approach, elastic net provides a balanced solution between these two techniques, resulting in more stable and relevant regression coefficient estimates. The steps for elastic net regression analysis are as follows [3].

- a. Selecting the optimal λ value through cross-validation with α value of $0 \leq \alpha \leq 1$.
- b. Performing ridge regression analysis with the optimal λ value using the formula

$$\hat{\beta}_{EN} = \arg \min \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda (a \sum_{j=1}^p |\beta_j| + (1 - a) \sum_{j=1}^p \beta_j^2) \right\} \quad (2.8)$$

2.7 Performance Evaluation

The selection of the best model can be determined using a model goodness of fit test, which employs R-squared (R^2), also known as the coefficient of determination. The coefficient of determination is an evaluation metric that measures how well a model fits the observed data and quantifies the amount of variation in the data that can be explained by the model. R^2 is calculated by dividing the variation explained by the model by the total variation in the data. The R^2 value ranges from 0 to 1, with higher values indicating that the model is better at explaining the data. To assess the goodness of the model, use the formula to calculate R^2 on the testing data as follows [4]:

$$R^2 = 1 - \frac{SSE}{SSR} \quad (2.9)$$

3. RESULT AND DISCUSSION

3.1 Parameter Estimation in the Elastic Net Mode

Estimating elastic net model parameters using coordinate descent optimization works by updating one parameter at each iteration while keeping the other parameters constant. The coordinate descent optimization method is used to estimate the parameters of the elastic net model. This method keeps other parameters fixed while updating one parameter at each iteration. In this way, objective functions that cannot be derived directly can be solved in optimization problems. The soft-thresholding operator is used to update the j th parameter at each iteration. This operator turns a parameter to zero if its value is less than a certain value, and reduces the absolute value of the parameter by that value. This operator applies regularization also known as LASSO regularization, which produces sparse models. Additionally, a cyclic coordinate descent rule is used to select which parameters to update and which to keep. This rule updates the parameters in turn, starting from the first parameter to the last, then back to the first parameter, and so on. Initially, the elastic net model is expressed in the following equation:

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda (a \sum_{j=1}^p |\beta_j| + (1 - a) \sum_{j=1}^p \beta_j^2) \quad (3.1)$$

Then, the partial derivative of equation (3.1) with respect to β_k with $k = 1, 2, \dots, p$ as follows:

$$\begin{aligned} \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k} &= \frac{\partial \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \left(a \sum_{j=1}^p |\beta_j| + (1-a) \sum_{j=1}^p \beta_j^2 \right) \right\}}{\partial \beta_k} \\ &= -2 \left(\sum_{i=1}^n x_{ik} (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij}) \right) + \lambda a (\text{sgn } \beta_k) + 2\lambda(1-a)\beta_k \end{aligned} \quad (3.2)$$

To determine the value of β_k , we can use equation (3.2) by equating the value to zero.

$$\begin{aligned} -2 \left(\sum_{i=1}^n x_{ik} (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij}) \right) + \lambda a (\text{sgn } \beta_k) + 2\lambda(1-a)\beta_k &= 0 \\ \hat{\beta}_k &= \frac{2 \sum_{i=1}^n x_{ik} (y_i - \beta_0 - \sum_{l \neq k}^p \beta_l x_{il}) - \lambda a (\text{sgn } \beta_k)}{2 \left(\sum_{i=1}^n x_{ik} (x_{ik}) \right) + 2\lambda(1-a)} \end{aligned} \quad (3.3)$$

meanwhile, for β_k with $k = 0$ as follows:

$$\begin{aligned} \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k} &= \frac{\partial \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \left(a \sum_{j=1}^p |\beta_j| + (1-a) \sum_{j=1}^p \beta_j^2 \right) \right\}}{\partial \beta_k} \\ &= - \left(\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij}) \right) \end{aligned} \quad (3.4)$$

To determine the value of β_k with $k = 0$, we can use the equation (3.4) by equating the value to zero:

$$\begin{aligned} - \left(\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij}) \right) &= 0 \\ \hat{\beta}_0 &= \frac{1}{n} \left(\sum_{i=1}^n (y_i - \sum_{j=1}^p \beta_j x_{ij}) \right) \end{aligned} \quad (3.5)$$

Equation (3.3) is used to update the β_k coefficient values with $k = 1, 2, \dots, p$ and Equation (3.5) is used to update the β_k coefficient values with $k = 0$ in the elastic net regression model using the coordinate descent method. This equation can be used iteratively until it reaches convergence.

3.2 Identification of Multicollinearity

The multicollinearity test aims to check whether there is a high correlation between the independent variables in the regression model. In a good regression model, the independent variables should not be closely related to each other. Test multicollinearity, we use the Variance Inflation Factor (VIF). VIF measures the extent to which the variance of a variable is influenced by other variables. The general criterion is if $VIF < 10$, then there is no multicollinearity problem. However, if $VIF > 10$, then there is an indication of a multicollinearity problem. This test is important to ensure accurate parameter estimation results and reduce bias in the regression model. Using the equation (2.4) The results of the multicollinearity test can be seen in Table (3.1) below.

Table 3.1. Multicollinearity Test Table

Independent variable	VIF
(X_1)	1.7910
(X_2)	30.2260
(X_3)	1.6428
(X_4)	3.1456
(X_5)	33.0336

From Table (3.1), it can be seen that there are VIF values that exceed 10, especially in the infant health service variable (X_2) and the number of mothers who received medical assistance during childbirth (X_5). This shows that there is a multicollinearity problem in the IMR data in South Sulawesi Province in 2019-2020 on the independent variable.

3.3 Cross Validation

In the cross validation stage, the dataset is divided into two subsets, namely the training data (80%) and the testing data (20%). This process adopts the k-fold cross validation method with a value of k equal to 5. Next, the subsequent step is to find the optimal values for α and λ , which are the penalty parameters in the ridge, LASSO, and elastic net models. The Determination of the optimal value is based on the criterion of minimizing the Root Mean Square Error (RMSE). By using equation (2.5), the cross validation results in table (3.2) can be obtained as follows:

Table 3.2. Cross validation results

Method	α	λ	RMSE
Ridge	0	1	3.4039
	0.1	1	3.3977
	0.2	0.7810	3.4099
	0.3	0.6120	3.4227
	0.4	0.5050	3.4337
Elastic net	0.5	0.4320	3.4393
	0.6	0.3810	3.4424
	0.7	0.3390	3.4433
	0.8	0.3040	3.4439
	0.9	0.2770	3.4441
LASSO	1	0.2520	3.4444

Based on the results presented in Table (3.2), we can find that in the ridge regression model, the optimal value for the penalty parameter λ is 1, with the smallest RMSE value being 3.4039. Meanwhile, in the LASSO regression model, we also obtain the lowest RMSE value of 3.4444, with the corresponding optimal λ value being 0.2520. On the other hand, in the Elastic net analysis, the lowest RMSE value is 3.3977. Therefore, in the elastic net regression, we will use the optimal α value of 0.1, while the found optimal λ value is 1.

3.4 Data Modeling Using Regularization Methods

In applying regularization methods in data analysis using the coordinate descent optimization technique based on equations (2.6), (2.7), and (2.8), the estimation of regression coefficient parameters for ridge, LASSO, and elastic net is conducted with the assistance of R-Studio software using the glmnet package. By utilizing the optimal α and λ values obtained in the cross validation table, the comparison of Ridge, LASSO, and elastic net regression coefficients can be observed as follows:

Table 3.3. Estimation of regression coefficient parameters

Variable	Ridge	LASSO	Elastic net
β_0	7.986544	7.494443	7.986544
(X_1)	-0.000079	-0.000024	-0.000079
(X_2)	-0.000125	0	-0.000125
(X_3)	-0.000088	-0.000218	-0.000088
(X_4)	0.000969	0.000934	0.000969

$$\begin{array}{cccc} (X_5) & -0.000116 & 0 & -0.000116 \end{array}$$

Based on the estimation results presented in Table (3.3), the regularization regression model using coordinate descent optimization is as follows:

a. Ridge Regression Model:

$$\hat{y}_i = 7.364082 - 0.000043X_1 - 0.000076X_2 - 0.00011X_3 + 0.003122X_4 - 0.000083X_5 \quad (3.6)$$

b. LASSO Regression Model:

$$\hat{y}_i = 7.494443 - 0.000024X_1 - 0.000218X_3 + 0.000934X_4 \quad (3.7)$$

c. Elastic net Regression Model:

$$\hat{y}_i = 7.361615 - 0.000038X_1 - 0.00006X_2 - 0.000113X_3 + 0.002177X_4 - 0.000066X_5 \quad (3.8)$$

from equations (3.5), (3.6), and (3.8) can be explained as follows:

1. Ridge Regression: From equation (3.5), it can be concluded that there is a positive influence on the level of public health in South Sulawesi Province from variables such as the number of babies who are exclusively breastfed, the increase in the number of baby health services, the number of vitamin A given to babies, and the number of mothers assisted in giving birth. medical in each district/city in South Sulawesi Province. However, for the variable number of low birth weight babies in each district/city in South Sulawesi Province, there is a negative influence on the level of public health in South Sulawesi Province.
2. LASSO Regression: From equation (3.6), there are only 3 variables that have an influence, namely the number of babies given exclusive breast milk, the amount of vitamin A given to babies, and the number of low birth weight babies. The variable number of low birth weight babies has a negative effect on the level of public health in South Sulawesi Province. and on the variable number of babies given exclusive breast milk, the amount of vitamin A given to babies has a positive effect.
3. Elastic net Regression: From equation (3.8), it can be concluded that there is a positive influence on the level of public health in South Sulawesi Province in 2019 from variables such as the number of babies who are exclusively breastfed, the increase in the number of baby health services, the number of vitamin A given to babies, and the number of mothers assisted in giving birth. medical in each district/city in South Sulawesi Province. However, there is a negative influence on the level of public health in South Sulawesi Province for the variable number of low birth weight babies in each district/city in South Sulawesi Province.

3.5 Goodness of fit of the regression model

The goodness of a regression model can be assessed using the R^2 value on the testing data. R-squared measures the extent to which the model can explain the variation in the dependent variable. A higher R-squared value indicates a better ability of the model to explain the variation in the data. Based on Equation (2.9), the comparison of R^2 values in Ridge, LASSO, and Elastic net regression can be observed in Table (3.4).

Table 3.4. coefficient of determination value

Method	R^2
Ridge	0.5411
LASSO	0.5818
Elastic net	0.6081

The results presented in Table (3.4) indicate that the Elastic-Net method is the best in assessing the goodness of fit for regression models utilizing coordinate descent optimization. When compared to other methods, elastic net has the highest R^2 value at 0.6081, followed by lasso with a value of 0.5818, and ridge with the lowest value of 0.5411. So it can be concluded that the elastic net method using coordinate descent optimization is the best method for overcoming extreme multicollinearity in infant mortality rate data in South Sulawesi 2019-2020.

4. CONCLUSION

The estimation results for infant mortality data in South Sulawesi Province using three methods, namely ridge regression, LASSO, and Elastic net, have been examined. By considering the goodness of fit indicator using the R^2 value, it can be concluded that Elastic net regression has the highest value compared to ridge and LASSO regression. Therefore, it can be concluded that there are five independent variables that influence infant mortality rates in South Sulawesi Province, namely exclusive breastfeeding, infant healthcare services, vitamin A supplementation, the number of low birth weight infants, and maternal conditions during pregnancy and childbirth. These variables have a positive impact on the public health level in South Sulawesi Province, except for the number of low birth weight infants in each district/city in South Sulawesi Province, which has a negative impact.

REFERENCES

- [1] Altalbany, S., 2021. Evaluation of Ridge, Elastic Net and Lasso Regression Methods in Precedence of Multicollinearity Problem: A Simulation Study. *Journal of Applied Economics and Business Studies*, Vol. 5, No. 1, 131–142.
- [2] Araveeporn, A., 2021. The higher-order of adaptive lasso and elastic net methods for classification on high dimensional data. *Mathematics*, Vol.9 No. 10, 1–14.
- [3] Fanny, R., Djuraidah, A., & Alamudi, A., 2018. Pendugaan Produktivitas Bagan Perahu dengan Regresi Gulud, LASSO dan Elastic-net. *Xplore: Journal of Statistics*, Vol. 2, No 2, 7–14.
- [4] Gujarati, D. N., 2021. Basic econometrics. In *Prentice Hall*.
- [5] Gupta, V. (2019). Evaluating the Accuracy of Valuation Multiples on Indian Firms Using Regularization Techniques of Penalized Regression. *Theoretical Economics Letters*, Vol. 9, No. 01, 180–209.
- [6] Handayani, A., & Wachidah, L., 2022. Metode Regresi Elastic-Net untuk Mengatasi Masalah Multikolinearitas pada Kasus Tingkat Pengangguran Terbuka di Provinsi Jawa Barat. *Bandung Conference Series: Statistics*, Vol. 2, No. 2, 459–465.
- [7] Ikhsan, E., & Ratu, N. Y., 2021. Small Area Estimation untuk Estimasi Angka Kematian Bayi di Indonesia menggunakan Metode Empirical Best Linear Unbiased Prediction. *Jurnal Biometrika Dan Kependudukan*, Vol. 10, No 2, 171–180.
- [8] Kusuma, G. W., & Wulansari, I. Y., 2019. Analisis Kemiskinan Dan Kerentanan Kemiskinan Dengan Regresi Ridge, Lasso, Dan Elastic-Net Di Provinsi Jawa Tengah Tahun 2017. *Seminar Nasional Official Statistics*, Vol. 1, No. 1, 503–513.
- [9] Mayapada, R., Tinungki, G. M., & Sunusi, N., 2018. Penerapan Sparse Principal Component Analysis dalam Menghasilkan Matriks Loading yang Sparse. *Jurnal Matematika Statistika Dan Komputasi*, Vol. 15, No. 2, 44.

- [10] Ningsih, S., & Dukalang, H. H., 2019. Penerapan Metode Suksesif Interval pada Analisis Regresi Linier Berganda. *Jambura Journal of Mathematics*, Vol. 1, No. 1, 43–53.
- [11] Palupi, W. R., & Rizki, L. K., 2020. Analysis of Factor Affecting Infant Mortality Rate (Imr) in East Java Using Multiple Linear Regression. *Jurnal Biometrika Dan Kependudukan*, Vol. 9, No. 1, 69–76.
- [12] Sartika, I., Naomi Nessyana, D., & Nurfitri, I., 2020. Analisis Regresi Dengan Metode Least Absolute Shrinkage and Selection Operator (Lasso) Dalam Mengatasi Multikolinearitas. *Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya*, Vol. 9, No. 1, 31–38.
- [13] Sonawane, M. A., & Tayade, A. Y., 2021. Predictive Model on Determinants of Child Mortality Using Multiple Regression Analysis. *Journal of Physics: Conference Series*, Vol. 1, No. 1, 1–6.
- [14] Srivastava, S., Upadhyay, S. K., Chauhan, S., & Alagarajan, M., 2021. Preceding child survival status and its effect on infant and child mortality in India: An evidence from National Family Health Survey 2015–16. *BMC Public Health*, Vol. 21, No. 1, 1–12.
- [15] Wicaksono, F., Susanto, B., & Farida, F., 2022. The Effect of Environmental Performance, Environmental Costs, and Environmental Disclosure on Profitability of Manufacturing Companies Listed in Indonesia Stock Exchange 2016 - 2020. *Urecol Journal. Part B: Economics and Business*, Vol. 2, No. 1, 32–42.
- [16] Wunderlich, K., & Thompson, E., 2018. Estimating Health Care Costs among Fragile and Conflict Affected States: An Elastic Net-Risk Measures Approach. *International Journal of Public Health Science (IJPHS)*, Vol. 7, No. 3, 175–185.
- [17] Zou, H., & Hastie, T., 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, Vol. 67, No. 2, 301–320.