Modelling The Recovery of Health Conditions for Pulmonary Tuberculosis Patients at Dr. Soegiri Hospital Lamongan in 2018 Using Binary Logistic Regression

S P Wulandari, D Susilaningrum and P D Viyanti

Business Statistics Department, Institut Teknologi Sepuluh Nopember, Sukolilo, Surabaya, Indonesia

Priska191@gmail.com

Abstract. East Java province occupied the second place with pulmonary tuberculosis cases in Indonesia. Lamongan is one of the regencies of East Java has a number of pulmonary tuberculosis treatment success that has not met the target, this means that there are still many pulmonary tuberculosis patients who do not recover after undergoing treatment or even died. Therefore conducted a study using a binary logistic regression method to find out the factors that affect recovery pulmonary tuberculosis sufferer's health condition in the Provincial Hospital Dr. Soegiri Lamongan so that people can avoid the factors that significantly affect the recovery of the health condition of pulmonary tuberculosis disease sufferers. The results showed that nutritional status, presence of complications, when influential significantly and long against the recovery of pulmonary tuberculosis patients condition.

1. Introduction

Pulmonary tuberculosis is an inflammation of the pulmonary parenchyma caused by Mycobacterium tuberculosis which is an acid-resistant aerobic bacilli that is transmitted by air (airbone)[3]. Indonesia occupied the second of five countries with the highest incidence of tuberculosis case in the world in 2016. East Java Province occupies second with the case of the most pulmonary tuberculosis in Indonesia. Based on the health profile of East Java Province in 2016, the figure of tuberculosis case findings of 23183 patients. Meanwhile, in Lamongan Regency in 2016 recorded 1901 people suffer from pulmonary tuberculosis. Based on the health profile of Lamongan Regency in 2016, the success rate of pulmonary tuberculosis is 89.36% which means that it has not fulfilled the target of the Ministry of Health by 90%. In the years 2018, Dr. Soegiri Lamongan hospital had 727 patients suffering from pulmonary tuberculosis and 159 of whom were not recovered after undergoing treatment. Many pulmonary tuberculosis patients who do not recover in Lamongan district are certainly caused by several factors. Therefore, this research is done to obtain a recovery model of the health condition of the pulmonary tuberculosis patients so that the factors that affect the recovery of health conditions of the pulmonary tuberculosis patients at Dr. Soegiri Lamongan hospital.

The response variables on this research are the recovery of health conditions of patients who are categorized into two namely recovered and not recovered so that the method used to analyze is a method of binary logistic regression. Binary logistic regression is a method of data analysis used to search for relationships between a binary (Y) or dichotomous variable response (X) that is a category or continuous predictor[4]. This research restricts to pulmonary tuberculosis patients at Dr. Soegiri...
Lamongan hospital in 2018 years of age and the hospitalization of Dr. Soegiri Lamongan in the condition of recovering or not recovering where it is not recovered in this case means that the patient asks to leave the hospital Before the increase or the patient is referred to another hospital or a deceased patient.

2. Review of The Library

2.1. Independence Test

The independence test is used to determine the relationship between two variables, namely the predictor variable with the response variable[1]. The independence test is conducted using the Pearson Chi-square test with the hypothesis as follows

H₀: No relationship between response variable with predictor variables
H₁: There is a relationship between response variable with predictor variables

The test statistics used are in accordance with equation 1 following

$$\chi^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$  \hspace{1cm} (1)

which

$$e_{ij} = \frac{n_{i.}n_{.j}}{n_{..}}$$  \hspace{1cm} (2)

Description :

$n_{ij}$: Frequencies for the I-row observation and the J-column

e$_{ij}$: Expectation value in the I-row observation and the J-column

If a significant level is $\alpha$ then H₀ is rejected if $\chi^2 > \chi^2_{(r-1),(c-1)}$ or $P$-value < $\alpha$

2.2. Binary Logistic Regression

Logistic regression is a regression analysis used to describe the relationship between nominal or ordinal-scale response variables with a continuous set of predictor variables or categorics [1]. Binary logistic regression is a method of data analysis used to search the relationship between a binary (dichotomus) response variable with a category or continuous predictor variable [4]. The result of a Dichotomus variable response has two criteria, i.e. $y = 1$ represents the probability of success with probability and $y = 0$ represents the possibility of failing with the probability in which the response variable follows Bernoulli's distribution for each single observation. Logistic regression models used are as follows

$$\pi(x) = \frac{\exp(g(x))}{1 + \exp(g(x))}$$  \hspace{1cm} (3)

2.2.1. Logistic regression parameters estimation. The general method of estimation of parameters in logistic regression is Maximum Likelihood Estimation (MLE). The MLE method estimates parameters by maximizing the likelihood function and having the requisite of a data should follow a specific distribution [4]. In binary logistical regression, each observation follows the distribution Bernoulli so that its likelihood function can be determined.

2.2.2. Simultaneous testing of parameters. Simultaneous testing of parameters used to define variables that significantly affect the response variables collectively[4]. The hypothesis used is given as follows

H₀: $\beta_1 = \beta_2 = \ldots = \beta_p = 0$

H₁: At least one of $\beta_m \neq 0$; $m = 1, 2, ..., p$ (without $\beta_0$)
Test statistics:

\[
G = -2 \ln \left( \frac{n_1}{n} \left( \frac{n_0}{n} \right)^{n_0} \right) \sum_{i=1}^{g} \hat{y}_i (1 - \hat{y}_i)^{1-y_i})
\]  
(4)

Description:

\( n_1 \): number of observations with category \( y=1 \)
\( n_0 \): number of observations with category \( y=0 \)
\( n \): number of observations
\( p \): number of parameters

\( G \) test statistics follow a chi-square distribution with a free degree is \( p \), if a significant level is \( \alpha \) then \( H_0 \) is rejected if the test value of \( G \) more than \( \chi^2(\alpha, p) \) or \( p \)-value is less than \( \alpha \)

2.2.3. Partially testing of parameters. Partially testing of parameters aims to know the parameters of which predictor variables have a significant effect on the response variables. The hypothesis used is given as follows

\( H_0: \beta_m = 0 \)
\( H_1: \beta_m \neq 0; m = 1, 2, ...., p \) (without \( \beta_0 \))

Test statistics:

\[
W^2 = \frac{\hat{\beta}_m^2 \text{ Se}(\hat{\beta}_m)}{\text{Se}(\hat{\beta}_m)^2}
\]  
(5)

Description:

\( \text{Se}(\hat{\beta}_m) \): Default parameter estimate error.
\( \hat{\beta}_m \): Parameter

\( W \) test statistics follow a chi-square, if a significant level is \( \alpha \) then \( H_0 \) is rejected if the test value of \( W^2 \) more than \( \chi^2(\alpha, p) \) or \( p \)-value is less than \( \alpha \)

2.2.4. Model Conformance Test. This test was conducted to test whether the model produced based on the multivariate/simultaneous logistics regression was feasible. This test uses Hosmer and Lemeshow test statistics [4] with hypotheses used as follows.

\( H_0 \): Appropriate models (no significant difference between observations and possible model prediction results)
\( H_1 \): Inappropriate models (no significant difference between observations and possible model prediction results)

Test statistics:

\[
\hat{c} = \sum_{i=1}^{g} \left( \frac{O_k - n_k \pi_k}{n_k \pi_k (1 - \pi_k)} \right)
\]  
(6)

Description:

\( O_k \): Observations in the \( k \) group
\( \pi_k \): Average estimated probability \( \left( \sum_{i=1}^{m} \frac{\hat{y}_i}{n_i} \right)/n_k \)
\( g \): Number of groups (combination of categories in simultaneous models)
\( n_k \): Number of observations in the \( k \) group
If a significant level is $\alpha$ then $H_0$ is rejected if $\hat{C} > \chi^2_{(n, \alpha, 2)}$.

2.2.5. Odds Ratio. Odds Ratio (OR) is one measure of the risk level used in interpreting the predictor variable coefficient. Odds ratios show a comparison of the chance of an incident with an probability not to appear. The odds ratio can be obtained from logistic regression values for predictor variables that are categorized 0 or 1 [4]. The odds ratio can be formulated as follows

$$OR = \frac{\pi(1)}{\pi(0)} = \frac{\exp(\beta_1 + \beta_2)}{\exp(\beta_1)} = \exp(\beta_2)$$ (7)

If OR = 1, then there is no link between the predictor variable with the response variable. If OR < 1, then there is a negative relationship between the predictor variable and the response variable on each x value change. If OR > 1, then there is a positive relationship between the predictor variable with the response variable on each x value change [4].

2.2.6. Accuracy of The Classification. Evaluation of the accuracy of the classification is an evaluation that sees the chance of classification errors based on criteria or size used [5]. Calculation of accuracy classification is as follows.

Table 1. Accuracy of The Classification

| Observation | $\pi_1$ | $\pi_2$ | Total |
|-------------|--------|--------|-------|
| $\pi_1$    | $n_{1C}$ | $n_{1M} = n_1 - n_{1C}$ | $n_1$ |
| $\pi_2$    | $n_{2M} = n_2 - n_{2C}$ | $n_{2C}$ | $n_2$ |

Description:

- $n_{1C}$: Values of the correct $\pi_1$ object are classified as $\pi_1$ object
- $n_{1M}$: Values of the $\pi_1$ object are incorrect classified as $\pi_2$ object
- $n_{2C}$: Values of the correct $\pi_2$ object are classified as $\pi_2$ object
- $n_{2M}$: Values of the $\pi_2$ object are incorrect classified as $\pi_1$ object

So that it can be formulated as follows

$$\text{Accuracy of The Classification} = \frac{n_{1C} + n_{2C}}{n_1 + n_2}$$ (8)

2.3. Tuberculosis

Tuberculosis is an infection of the disease caused by Mycobacterium tuberculosis which is an acid-resistant aerobic bacilli that is transmitted by air (airbone). Tuberculosis infection is obtained through a fairly small inhalation of germ particles. Before pulmonary infection can occur, the organism that is inhalation must first fight against pulmonary defense mechanisms and pulmonary tissue [2]. Pulmonary tuberculosis is an inflammatory disease of the pulmonary parenchyma due to the infection with Mycobacterium tuberculosis. Pulmonary tuberculosis covers 8% of the overall incidence of tuberculosis disease, while 20% is the most exbiotic tuberculosis[6].

3. Research Metodology

3.1. Data Source

Data used in this research is secondary data in the form of medical record in Dr. Soegiri Lamongan hospital is the health condition data of the pulmonary tuberculosis patients and the factors influencing it from January 1st 2018 until December 31st 2018 As many as 100 medical record data that has been prepared by the staff of Dr. Soegiri Lamongan hospital from a population of 727 data.
3.2. Variable Research
The research variables used in this study consist of response variables and predictor variables. The explanation of each variable is as follows.

3.2.1. Response Variable. The response variable \( (Y) \) in this study is the condition of a pulmonary tuberculosis patient in Dr. Soegiri Lamongan hospital.

| Table 2. Response Variable |
|----------------------------|
| Variable | Description | Category | Scale |
| Y | The condition of pulmonary tuberculosis patient | 0 : Not Recovered | Nominal |
| | | 1 : Recovered | |

3.2.2. Predictor Variable. The Predictor variables \( (X) \) used in this study are shown in the following table.

| Table 3. Predictor Variable |
|----------------------------|
| Variable | Category | Scale |
| Age \( (X_1) \) | 0 : Non productive (<15 years or >50 years) | Nominal |
| | 1 : Productive (15-50 years) | |
| Last Education \( (X_2) \) | 0 : ≤ Elementary School | Ordinal |
| | 1 : Junior High School | |
| | 2 : ≥ Senior High School | |
| Smoking Status \( (X_3) \) | 0 : Yes | Nominal |
| | 1 : No | |
| Nutritional Status \( (X_4) \) | 0 : Less | Nominal |
| | 1 : Quite | |
| Complication \( (X_5) \) | 0 : Existing | Nominal |
| | 1 : Nothing | |
| Cough History \( (X_6) \) | 0 : ≥ 6 months | Nominal |
| | 1 : < 6 months | |
| Fever \( (X_7) \) | 0 : Yes | Nominal |
| | 1 : No | |
| Shortness of Breath \( (X_8) \) | 0 : Yes | Nominal |
| | 1 : No | |
| Chest Pain \( (X_9) \) | 0 : Severe pain | Ordinal |
| | 1 : Mild pain | |
| | 2 : No pain | |
| Length of Stay in Hospital \( (X_{10}) \) | 0 : < 6 days | Nominal |
| | 1 : ≥ 6 days | |

3.3. Data Analysis Steps
The analysis steps done in this study are as follows.
1. Performs the independence test between the response variables \( (Y) \) with predictor variables \( (X) \)
2. Conducting a binary logistic regression analysis to model the alleged factors affecting the recovery of the condition of pulmonary tuberculosis patients with the following stages of analysis.
   a. Perform estimation of binary logistic regression parameters.
   b. Test the significance of the parameters simultaneously.
   c. Test the significance of the parameters partially.
   d. Model formation and interpret the model formed.
   e. Calculates and interprets the odds ratio.
   f. Conduct a model conformance test
   g. Calculating the accuracy of the model classification
3. Draw conclusions from the results of analysis

4. Analysis Result

4.1 Independence Test Analysis
The independence test is used to determine the presence or absence of the relationship between the recovery of the condition of the pulmonary tuberculosis patients with the allegedly affected factors. The hypotheses used are as follows.
H$_0$: There is no relationship between the recovery of the condition of pulmonary tuberculosis patients with the suspected variable influencing them.
H$_1$: There is a relationship between the recovery of the condition of pulmonary tuberculosis patients with the suspected variable influencing them.
Significant levels : $\alpha = 0.05$

Rejected area : $H_0$ rejected if $\chi^2 > \chi^2_{0.05, df}$ or p-value < $\alpha$

Test statistical results are shown in the table below.

| Variable                      | $\chi^2$ | $\chi^2_{0.05, df}$ | Df | P-value |
|-------------------------------|----------|----------------------|----|---------|
| Age ($X_1$)                   | 0.069    | 3.841                | 1  | 0.792   |
| Last Education ($X_2$)        | 0.392    | 5.991                | 2  | 0.825   |
| Smoking Status ($X_3$)        | 0.126    | 3.841                | 1  | 0.723   |
| Nutritional Status ($X_4$)    | 7.310    | 3.841                | 1  | 0.007*  |
| Complication ($X_5$)          | 12.077   | 3.841                | 1  | 0.001*  |
| Cough History ($X_6$)         | 5.315    | 3.841                | 1  | 0.031*  |
| Fever ($X_7$)                 | 0.236    | 3.841                | 1  | 0.638   |
| Shortness of Breath ($X_8$)   | 0.152    | 3.841                | 1  | 0.696   |
| Chest Pain ($X_9$)            | 4.531    | 5.991                | 2  | 0.091   |
| Length of Stay in Hospital ($X_{10}$) | 9.013 | 3.841                | 1  | 0.003*  |

*) significant with $\alpha = 0.05$

The table shows that the value of nutritional status, complications, cough history, and length of stay is greater than $\chi^2_{0.05, df}$, reinforced with a smaller p-value of a significant amount of 0.05 resulting in a decision to reject $H_0$. So it can be concluded that there is a relationship between the recovery of the condition of pulmonary tuberculosis patients with nutritional status, complications, cough history, and the length of patient stay.

4.2 Binary Logistic Regression Analysis
Binary logistic regression is a method of data analysis used to find the relationship between a dichotomous response variable with a polichotomous predictor variable. Here is the result of binary logistic regression analysis to know the factors that affect the recovery of the condition of pulmonary tuberculosis patients.
4.2.1 Simultaneous testing of Parameter. Simultaneous testing was conducted to determine the influence of predictor variables on the recovery of health conditions of pulmonary tuberculosis patients simultaneously. The hypotheses used are as follows.

\( H_0 : \beta_1 = \beta_2 = \beta_3 = \ldots = \beta_{12} = 0 \) (Predictor variables have no effect on the response variables)

\( H_1 : \beta_m \neq 0, m=1,2,3,\ldots,12 \) (There are at least 1 predictor variable that affects the response variables)

Significant level : \( \alpha = 0.05 \)

Rejected area : \( H_0 \) rejected if \( \chi^2 > \chi^2_{\alpha, df} \) or \( P\text{-value} < \alpha \)

Test statistical results are shown in the table below.

**Table 5. Simultaneous Testing**

| \( \chi^2 \) | \( \chi^2_{0.05} \) | Df | P-value |
|---------------|----------------|----|---------|
| 37.016        | 21.026         | 12 | 0.000   |

The table shows that the value of \( \chi^2 = 37.016 \) is greater than \( \chi^2_{0.05,12} = 21.026 \) and P-value of 0.000 is less than the significant level of 0.05 so that the decision is rejected \( H_0 \) which means that there are one or more predictor variables affect the response variables.

4.2.2 Partially testing of Parameter. Simultaneous testing carried out that there was one or more predictor variables that had an effect on the response variables, therefore a partial test is performed to determine which variables are significantly influential to the recovery of pulmonary tuberculosis patients with the following hypotheses.

\( H_0 : \beta_m = 0 \)

\( H_1 : \beta_m \neq 0, m=1,2,3,\ldots,12 \)

Significant level : \( \alpha = 0.05 \)

Rejected area : \( H_0 \) rejected if \( W^2 > \chi^2_{\alpha} \) or \( P\text{-value} < \alpha \)

Test statistical results are shown in the table below.

**Table 6. Partially Testing**

| Variable                  | \( W^2 \) | \( \chi^2_{0.05} \) | P-value |
|---------------------------|----------|----------------|--------|
| Age (1)                   | 0.009    | 3.841          | 0.926  |
| Last Education (1)        | 0.112    | 3.841          | 0.738  |
| Last Education (2)        | 1.048    | 3.841          | 0.306  |
| Smoking Status (1)        | 0.652    | 3.841          | 0.419  |
| Nutritional Status (1)    | 6.303    | 3.841          | 0.012* |
| Complication (1)          | 8.349    | 3.841          | 0.004* |
| Cough History (1)         | 0.007    | 3.841          | 0.935  |
| Fever (1)                 | 0.775    | 3.841          | 0.379  |
| Shortness of Breath (1)   | 0.459    | 3.841          | 0.498  |
| Chest Pain (1)            | 2.354    | 3.841          | 0.125  |
| Chest Pain (2)            | 2.395    | 3.841          | 0.122  |
| Length of Stay in Hospital (1) | 8.560    | 3.841          | 0.003* |

*) significant with \( \alpha = 0.05 \)

The table shows that the \( W^2 \) value of the nutritional status, complications, and length of stay was greater than 3.841 as well as a smaller p-value of 0.05. This indicates that the status of nutrition,
complications, and length of stay has significant effect on restoring the condition of pulmonary tuberculosis patients.

4.2.3 Binary Logistik Regression Model Formation. Forming a binary logistic regression model on the recovery of the condition of the pulmonary tuberculosis patients carried out by inserting a significant predictor variable that affects the condition of pulmonary tuberculosis patients. The test results are significance parameters simultaneously using the following hypotheses

\[ H_0: \beta_1 = \beta_2 = \beta_3 = 0 \] (Predictor variables have no effect on the response variable)

\[ H_1: \beta_m \neq 0, m=1,2,3 \] (There are at least on predictor variable that affect response variable)

Significant level: \( \alpha = 0.05 \)

Rejected area: \( H_0 \) rejected if \( \chi^2 > \chi^2_{\alpha, df} \) or \( P\text{-value} < \alpha \)

Test statistical results are shown in the table below.

| \( \chi^2 \) | \( \chi^2_{0.05, \alpha} \) | Df | P-value |
|--------------|------------------|----|---------|
| 29.457       | 7.814            | 3  | 0.000   |

The table indicates that the value of \( \chi^2 \) 29.457 is greater than \( \chi^2_{0.05, \alpha} \) 7.814 and P-value of 0.000 is less than the significant level of 0.05 so that the decision is rejected \( H_0 \) which means there is one or more of the nutritional status variables, Complications, and long stays that affect the recovery of the condition of pulmonary tuberculosis patients.

The test results are significance parameters partially using the following hypotheses

\[ H_0: \beta_m = 0 \]

\[ H_1: \beta_m \neq 0, m = 1,2,3 \]

Significant level: \( \alpha = 0.05 \)

Rejected area: \( H_0 \) rejected if \( W^2 > \chi^2_{\alpha, df} \) or \( P\text{-value} < \alpha \)

Test statistical results are shown in the table below.

| Variable                  | B     | \( W^2 \) | \( \chi^2_{0.05,1} \) | P-value |
|---------------------------|-------|-----------|----------------------|---------|
| Nutritional Status (1)    | -1.434| 5.424     | 3.841                | 0.020*  |
| Complication (1)          | -1.945| 9.354     | 3.841                | 0.002*  |
| Length of Stay in Hospital(1) | -1.857| 9.612     | 3.841                | 0.002*  |
| Constant                  | 3.877 | 26.469    | 3.841                | 0.000   |

*) significant with \( \alpha = 0.05 \)

The table shows that the \( W^2 \) value of the nutritional status, complications, and length of stay was greater than 3.841 as well as a smaller p-value of 0.05. This indicates that the status of nutrition, complications, and length of stay has significant effect on restoring the condition of pulmonary tuberculosis patients.

4.2.4 Binary Logistik Regression Logit Model. Based on simultaneous and partial testing results in model formation, the logit model in binary logistic regression analysis on the recovery of pulmonary tuberculosis patient conditions is as follows.

\[ \hat{g}(x) = 3.877 - 1.434X_4(l) - 1.945X_5(l) - 1.857X_{10}(l) \]

Based on the logit model, the value of probability gained if a pulmonary tuberculosis patient with less nutritional status, has complications, and is hospitalized for less than 6 days. Patients with the condition have the probability to recover 0.20442, while the probability to not recover is 0.79558 in
accordance with the following calculation where $\pi_1$ is the probability of pulmonary tuberculosis patients to recover and $\pi_2$ is the probability of pulmonary tuberculosis patients not recovered.

$$\pi_1 = \frac{\exp[g(x)]}{1 + \exp[g(x)]}$$

$$\pi_1 = \frac{\exp\{3.877 - 1.434X_{i1} - 1.945X_{i2} - 1.857X_{i3}\}}{1 + \exp\{3.877 - 1.434X_{i1} - 1.945X_{i2} - 1.857X_{i3}\}}$$

$$\pi_1 = \frac{\exp(3.877 - 1.434(1) - 1.945(1) - 1.857(1))}{\exp(-1.359)}$$

$$\pi_1 = 0.20442$$

$$\pi_2 = 1 - \pi_1$$

$$\pi_2 = 1 - 0.20442 = 0.79558$$

4.2.5  Odds Ratio. The odds ratio of each predictor variable entered in the binary logistic regression model is shown in the table below.

| Variable                   | Odds Ratio |
|----------------------------|------------|
| Nutritional Status (1)     | 0.238      |
| Complication (1)           | 0.143      |
| Length of Stay in Hospital(1) | 0.156 |

The table shows that a lung patient with less nutritional status has a chance to recover by 0.238 times compared to pulmonary tuberculosis patients with adequate nutritional status. If a pulmonary TB patient has complications, then the chance to recover is 0.143 times compared to pulmonary tuberculosis patients who have no complications. If pulmonary tuberculosis patients are hospitalized for less than 6 days, the patient's chance to recover is 0.156 times compared to patients who are hospitalized for more than or equal to 6 days.

4.2.6  The Suitability Test Model Recovery of Pulmonary Tuberculosis Patient. The model conformance test is used to determine whether the model has been acquired accordingly. Hypotheses used in the conformance test model recovery of pulmonary tuberculosis patients are as follows.

$H_0$: Model has been appropriate (no significant difference between observation results with possible model prediction results)

$H_1$: Inappropriate Model (there is a significant difference between observation results and possible model prediction results)

Significant level : $\alpha = 0.05$

Rejected area : $H_0$ rejected if $\chi^2 > \chi^2_{\alpha, df}$ or $P-value < \alpha$

Test statistical results are shown in the table below.

| $\chi^2$ | $\chi^2_{0.05, df}$ | Df | P-value |
|----------|----------------------|----|---------|
| 1.268    | 11.070               | 5  | 0.938   |

The table shows that the value of $\chi^2 = 1.268$ is less than $\chi^2_{0.05, df} = 11.070$ which is reinforced with P-value of 0.938 greater than the significant level of 0.05 so that the decision is failed to reject $H_0$. So it
can be concluded that the binary logistic regression model on the recovery of the condition of pulmonary tuberculosis patients already appropriate.

4.2.7 Accuracy of The Classification. Evaluation of the classification accuracy is an evaluation to see the chance of classification errors based on criteria or size used is Accuracy Total. Assessment result of exactness of recovery classification of pulmonary tuberculosis patients shown in the table below.

**Table 11. Accuracy of The Classification**

| Observation of The Patient’s Condition | Prediction of the Patient’s Condition | Percentage |
|---------------------------------------|--------------------------------------|------------|
| Not Recover                           | 7                                    | 18         | 28%        |
| Recovered                             | 2                                    | 73         | 97.3%      |
| Total Percentage                      |                                      |            | 80%        |

The table shows that the percentage of the accuracy of the classification of unrecovered pulmonary tuberculosis patients is 28% where from 25 unrecovered pulmonary TB patients, 7 patients were correctly classified as unrecovered patients, while 18 patients Others are classified incorrectly as recovering patients. The percentage of the correctness of the recovered patients was 97.3% which of the 75 patients recovered pulmonary tuberculosis, 73 patients were correctly classified as recovered patients while two other patients were classified incorrectly as patients Not recover. Based on this, the overall percentage of the classification is 80% which means that from 100 pulmonary tuberculosis patients, 80 patients are correctly classified.

5. Conclusions And Suggestions

5.1 Conclusions

Based on analysis and discussion, the conclusion is the status of nutrition, complications, cough history, and the length of stay has a relationship with the recovery of the condition of pulmonary tuberculosis patients. The logit model obtained is

\[ \hat{g}(x) = 3.877 - 1.434X_1 + 1.945X_4 - 1.857X_{10} \]

that based on the model, the factors that have a significant effect on the recovery of the condition of pulmonary tuberculosis patients is the Status of nutrition, complications, and length of stay.

5.2 Suggestions

Based on the conclusions obtained, the advice from this research can be addressed to the local government of Lamongan Regency and the party of Dr. Soegiri Lamongan as the policy determinant in tackling pulmonary tuberculosis disease and other diseases to Reducing the percentage of pulmonary tuberculosis patients who were not recovered when leaving the hospital because the results showed that there were complications of significant effect on the recovery of the patient's condition. In addition, the people of Lamongan Regency are also advised to keep the diet so that the nutrients have been achieved, and to conduct treatment early if they have prolonged cough because the results explain that both of these things have a significant relationship with the restoration of patient condition.

References

[1] Agresti A 2002 *Categorical Data Analysis*. New York: John Wiley & Sons, Inc.
[2] Asih N G, Effendy C 2004. *Keperawatan Medikal Bedah*. Jakarta: EGC.
[3] Djojodibroto R D 2007. *Respirologi*. Jakarta: EGC.
[4] Hosmer D W, Lemeshow S 2000. *Applied Logistic Regression*. New York: John Wiley & Sons, Inc.
[5] Johnson R, Winchern D 2007. *Applied Multivariate Statistical Analysis*. New Jersey: Pearson Education, Inc.

[6] Novianti R 2016. *Struktur dan Fungsi Perawatan Kesehatan Terhadap Ketahanan Kesehatan pada Keluarga dengan Resiko Penularan TB Paru*. Program Studi Keperawatan STIKES Ngudia Husada Madura.