EVALUATION OF CONTAINER TERMINAL SYSTEM PERFORMANCE IN TANJUNG PERAK SURABAYA PORT IN INCREASING CUSTOMER SATISFACTION USING BINARY LOGISTICS REGRESSION

Muhammad Raehan Khaeruddin[1*], Tri Achmadi[1]

[1] Department of Marine Transportation Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
E-mail: mraehank.35@gmail.com

Received: January 1, 2024 | **Accepted:** January 31, 2024 | **Published:** February 1, 2024
DOI: 810.12962/j27745449.v4i2.1058 | **Issue:** Volume 4 Number 2 2023 | **E-ISSN:** 2774-5449

ABSTRACT

In this sophisticated era, we can do almost anything, and we can get it easily. No exception with the necessities of life that we can get easily from one place and sent to us via transportation services, even goods with a large capacity. One of the transportation services that can efficiently deliver goods in large quantities is sea transportation via ships with a shipping system using containers. Tanjung Perak Port in Surabaya is one of the ports that provides container loading and unloading services. For this reason, in this study an evaluation of the performance of container terminals at the Port of Tanjung Perak Surabaya will be carried out by looking at the factors that influence customer satisfaction (Y), which consist of several predictor variables, namely price of goods (X1), weight of goods (X2) and Travel Time (X3) using the binary logistic regression method, which is a method used to model the response variable consisting of two categories and is appropriate for modeling data consisting of possible events with response variables consisting of two choice categories. Logistic regression will form a predictor/response variable, which is a linear combination of the independent variables. In this study, modeling was carried out to determine the relationship between whether the predictor variable affects the response variable and the extent to which this variable influences it. After knowing the variables that contribute to customer satisfaction, these variables can then be used as evaluation material, whether they need to be increased or reduced. Based on the results of the study, it was found that the variable weight of goods (X2) has a significant influence on customer satisfaction with a classification accuracy of 88.1%, so it must be optimized related to the container capacity that contains customer goods because, if this variable increases by 1 unit, it will increase the probability of customer satisfaction is 1.792 times and this increase is higher than the other variables.

Keyword: Binary Logistics Regression, Containers, Port of Tanjung Perak, Satisfaction Level

Introduction

Indonesia is a maritime country with an ocean area of 3,257,357 km2 based on the results of the International Law of the Sea Convention or the so-called United Nations Convention on the Law of the Sea (UNCLOS) on December 10, 1982, in Montego Bay, Jamaica [1]. For this reason, because the sea area is very wide and Indonesia is an archipelagic country, activities carried out between islands or regions, both activities in the economic, educational, social, and political fields, require means of transportation that connect them, one of the transportation alternatives used by the majority of people to stay connected. And can carry out inter-island or regional activities by using sea transportation or what is commonly called a ship. Ships can transport passengers and goods from one place to another through the port. The port is a water area that is protected against waves equipped with sea terminal facilities, which include a pier where ships can moor for loading and unloading of goods and passengers and a storage area where ships unload their cargo and warehouses where goods can be stored, in a longer time while waiting for de-livery to the destination or customer [2]. Tanjung Perak is one of the busiest ports in Indonesia, which is located in Surabaya City, East Java Province, which functions as a collector and distributor of goods to and from eastern Indonesia, including East Java province. Due

to its strategic location and sup-ported by potential hinterlands, Tanjung Perak port is also an Interinsular Shipping Center for Eastern Indonesia. Tanjung Perak Port is included in Pelindo III, which is responsible for managing ports in 7 provinces, namely East Java, Central Java, South Kalimantan, Central Kalimantan, Bali, NTB, and NTT [3]. This port is one of the factors that play an important role in eco-nomic progress in Eastern Indonesia, especially because shipping operations focus on Eastern Indonesia and its surroundings. Not only that, the flow of the economy is also growing because this port opens shipping for the export and import of goods using containers.

According to the Central Bureau of Statistics (BPS) for East Java province, the volume of goods from export and import activities carried out at the Port of Tanjung Perak Surabaya has increased every year, as in 2020 the volume of goods reached 17,598,146,661 Kg while in the following year, namely in 2021, it has increased to 18,870,972,615 or an increase of 7% in 2021 com-pared to the previous year [4]. As one of the busiest ports in Indonesia, Tanjung Perak Port always updates and improves infrastructure and service facilities at the port, both in terms of services for passengers and loading and un-loading of containers. However, even so, the success of improving the facilities provided can be measured by customer satisfaction from Tanjung Perak Port in Surabaya. As for this study, customer satisfaction is measured by three variables, namely, price of goods (X1), weight of goods (X2), and travel time (X3) using the binary logistic regression method. Binary logistic regression is a method used to see the relationship between the response variable (y), which is binary or dichotomous, and the predictor variable (x), which is polychotomous or ratio [5]. After analyzing customer satisfaction data at Tanjung Perak Port, there are several studies that use binary logistic regression, but the majority of studies only focus on community social problems; therefore, this research focuses on evaluating customer satisfaction at Tanjung Perak Port, Surabaya.

Methodology

Independence Test

In this test, we will investigate hypotheses that have to do with whether or not two random variables take their values independently or whether the value of one has a relation to the value of the other. Thus, the hypotheses will be expressed in words, not mathematical symbols [6]. The hypothesis used in the independence test is.

H0: There is no relationship between the observed variables

H1: There is a relationship between the observed variables

Mathematically, the independence test can be calculated using the following equation [6].

$$\chi^2 = \sum_{k=1}^{v}\sum_{m=1}^{w}\frac{(O_{km}-E_{km})^2}{E_{km}} \quad (1)$$

Where Okm is the actual observation data of the kth row of the m-th column while Ekm is the expected value of the k-throw of the m-th column with degrees of freedom and a test results, in conclusion, rejecting H0 if 〖χ^2〗 _count> 〖χ^2〗 _table or when p-value > α (0.05) which means that the assumption of independence has been met.

Numerical Modeling of Debonding Damage

In Logistic regression is an excellent tool for modeling relationships with outcomes that are not measured on a continuous scale (a key requirement for linear regression). Logistic regression is often leveraged to model the probability of observations belonging to different classes of a categorical outcome, and this type of modeling is known as classification [7]. The response variable is dichotomous qualitative data with a value of 1 (one) to indicate the occurrence of an event and a value of 0 (zero) to indicate the non-occurrence of an event [8]. The value of the predictor variable in binary logistic regression will be transformed into a probability with a log function. The binary logistic regression equation model used in this study is,

$$\pi(x) = \frac{exp(\beta_0+\beta_1 x_1+\cdots+\beta_n x_n)}{1+exp(\beta_0+\beta_1 x_1+\cdots+\beta_n x_n)} \quad (2)$$

Where p is the number of predictor variables and $\pi(x)$ is the probability of an event being successful (yes or satisfied or successful) and needs to be transformed into a logit form because it is a nonlinear function so that the relationship between predictor and response variables can be interpreted [5].

Simultaneous Test

Simultaneous testing uses the likelihood ratio to test the feasibility of the model obtained from the parameter estimation results. The simultaneous test aims to determine whether the predictor variables contained in the model have a significant effect on the whole or not [8] with hypothesis,

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_n = 0$

$H_1 :$ at least one $\beta_r \neq 0$; r = 1,2,⋯, n

Simultaneous or simultaneous test calculations mathematically can be seen in the following equation (3), which is a research test statistic.

$$G^2 = -2\,ln\left(\frac{\left(\frac{n_1}{n}\right)^{n_1}\left(\frac{n_0}{n}\right)^{n_0}}{\prod_{i=1}^{n}\hat{\pi}_i^{y_i}(1-\hat{\pi}_i)^{(1-y_i)}}\right) \quad (3)$$

Tests using the G2 test statistic are known to follow the chi-square distribution and reject H0 if the p-value < 0.05 which means that all predictor variables jointly affect the response variable [8].

Partial Test

A partial test using the Wald test has the aim of knowing the significance of the parameters on the predictor variables [8].

$H_0 : \beta_0 = 0$

$H_1 : \beta_r \neq 0$; r = 1,2,⋯, n

The test statistic, which is a mathematical calculation of the Wald test, is presented in the equation (4) bellow [8].

$$W = \frac{\hat{\beta}_r}{SE(\hat{\beta}_r)} \quad (4)$$

Where is the parameter estimate and $SE(\hat{\beta}_k)$ is an estimate of the standard error. H_0 is rejected if the p-value < α or $|W| > Z_{\frac{\alpha}{2}}$ or $W^2 > \chi^2{}_{\alpha}$.

Model Fit Test

The suitability test of the model used is the Hosmer and Lemeshow test, which aims to test the suitability of the binary logistic regression model using the following hypothesis [8].

H_0 : The resulting model is suitable

H_1 : The resulting model is not suitable

The equation used in the Hosmer and Lemeshow tests to test the suitability of the data is.

$$\hat{C} = \sum_{h=1}^{g}\frac{(O_h - n_h\bar{\pi}_h)^2}{n_h\bar{\pi}_h(1-n_h\bar{\pi}_h)} \quad (5)$$

With,

g = the number of group

O_h = the sum of the response variable values

$\bar{\pi}_h$ = the average of the probability estimates

n_h = the number of subjects in the h group

Conclusion reject H_0 if $\hat{C} > \chi^2{}_{\alpha}$ which means the result of prediction or estimation with data.

Classification Accuracy

The classification table explains how well the model grouping cases into two groups [8] and to find out whether the classification results we get are correct, we can use the calculation of the accuracy of the classification based on the Apparent Error Rate (APER) with the formula [9].

Classification Accuracy = $(1 - APER) \times 100\%$ (6)

$$= 1 - \frac{n_{12}+n_{21}}{n_{11}+n_{12}+n_{21}+n_{22}} \quad (7)$$

where,

n_{11} = number of subjects from Y_1 rightly classified in Y_1

n_{12} = number of subjects from Y_1 classified in Y_2

n_{21} = number of subjects from Y_2 classified in Y_1

n_{22} = number of subjects from Y_2 rightly classified in Y_2

Data Source

The data used in this study is primary data from a survey of 67 people regarding their satisfaction with the performance of the container system at Tanjung Perak Port in Surabaya with Goods Price (X_1), Weight of goods (X_2) and Travel Time (X_3). as the predictor variable and customer satisfaction as the response variable (Y).

Result and Discussion

Descriptif Statistics

Tanjung Perak Port in Surabaya is the second busiest port in Indonesia after Tanjung Priok Port in Jakarta and is a trading center for eastern Indonesia [10]. For this reason, before analyzing customer satisfaction further, it is necessary to know in advance the characteristics of each ratio scale predictor variable used in the study as Table 1.

Table 1. Descriptive Statistics of Research Variables

Variable	Mean	Variance	Min	Max
X_1	93582	46165762099	550000	1500000
X_2	6.701	4.303	2.000	10.000
X_3	6.299	3.637	2.000	10.000

Table 1 provides individual information regarding the descriptive statistics of each predictor variable used, namely price of goods (X_1), weight of goods (X_2) and travel time (X_3), where the table shows that the variable price of goods (X_1) has a variant that the largest among other variables which shows that the distribution of data for this variable is very diverse. Furthermore, for the response variable which is a categorical scale variable, the percentage of customer satisfaction will be seen based on the data obtained as

Figure 1 Based on Figure 1 it is known that the majority of customers who have used container loading and unloading services at Tanjung Perak Port are satisfied (category 1) with the services and facilities provided at Tanjung Perak Port with a percentage of 86.6% of the total population who filled out the questionnaire.

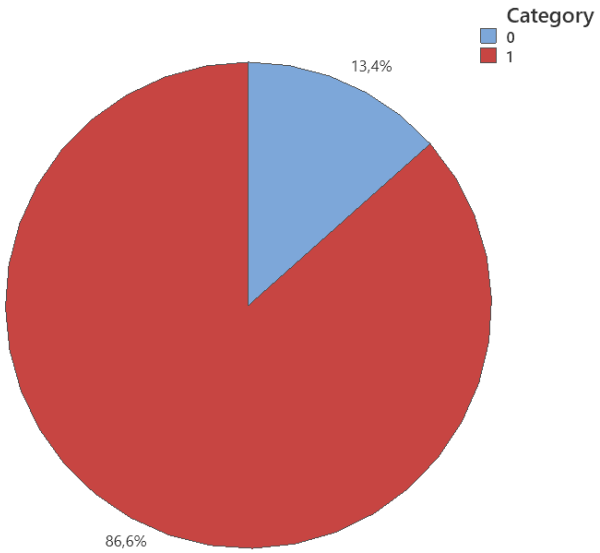

Figure. 1 Pie Chart of Response Variable

Independence Test

The independence test in this study used the chi-square test with the hypothesis that.

H_0 : There is no relationship between the observed variables

H_1 : There is a relationship between the observed variables

The significance level (alpha) used in this study was 0.05 with the results of the independence test for each research variable presented in Table 2. Based on Table 2, it is known that the p-value of each variable is greater than the significance level, which is 0.05, which can be concluded that it fails to reject H0, which means that there is no relationship between variables so that the assumption of independence has been fulfilled.

Table 2. Output Independence Test

Variable	P-value	Information
X_1	0.052	No Relationship
X_2	0.153	No Relationship
X_3	0.700	No Relationship

Simultaneous Test

The hypothesis used in this study adjusted for the number of predictor variables is,

$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

$H_1 :$ at least one $\beta_r \neq 0$; r = 1,2,3

The significance level used was 0.05 and the result was that the resulting G^2 test value was 8.603 with a p-value of 0.035 where the p-value is less than 0.05 so that it can be concluded that H_0 is rejected, which means that there is at least one variable that significantly influences customer satisfaction, so that it can be continued with a partial test.

Partial Test

The partial test is carried out by testing each variable and if the simultaneous test is obtained, the conclusion is reject H_0. Partial test is good to use on data where heteroscedasticity does not occur [11]. In this study, the hypothesis used.

$H_0 : \beta_0 = 0$

$H_1 : \beta_r \neq 0$; r = 1,2,3

The partial test results for each predictor variable for the response variable are shown in Table 3. It can be seen from the Table 3 that using a significance level of 0.05, there is only 1 variable that influences customer satisfaction for container loading and unloading service users at the Port of Tanjung Perak Surabaya, namely the variable weight of goods (X_2) with the resulting binary logistic regression model is,

$$\hat{g}(x) = 1.416 + 0.584 X_2 \quad (8)$$

Table 3. Output Partial Test

Variable	B	Wald	P-Value
X_1	0.000	1.969	0.161
X_2	0.584	7.301	0.007
X_3	-0.890	0.155	0.694
Constant	1.416	0.609	0.435

Model Fit Test

Goodness of fit is often used to assess how well a given probability distribution fits the data as well as how a statistical regression model fits the data [12]. Testing the suitability of the model can be done with the Hosmer and Lemeshow tests with the hypothesis used in the test being [13],

H_0 : The resulting model is appropriate

H_1 : The resulting model is not suitable

Based on the tests carried out, it was found that the value of the test statistic was 10.738 with a degree of freedom is 7 and a p-value is 0.150. If the resulting p-value is compared with the 0.05 significance level, it is known that the p-value > significance level (alpha), which means it fails to reject H_0. That is, the model produced in this study is in accordance with the actual

data and the estimated data using binary logistic regression.

Classification Accuracy

The model that has been obtained beforehand must first know the accuracy of the classification based on the criteria of the variable response used. The results of measuring the accuracy of the classification in this study can be seen in **Error! Reference source not found.**. Based on the results of the classification accuracy obtained, an overall classification accuracy of 88.1% was obtained. The resulting classification accuracy percentage is high enough, so this model can already estimate the observation data for further predictions.

Model Interpretation

Interpretation of the model can be carried out by looking at the odds ratio value generated in the modeling using binary logistic regression as shown in **Error! Reference source not found.**. Based on the odds ratio of each variable, it can be seen that if the price of a customer's goods increases by 1 unit, it will increase the probability of customer satisfaction by 1.000 times. As for the variable weight of goods, if it increases by one unit, it will increase the probability of customer satisfaction by 1.792 times and if the variable travel time increases by one unit, it will increase the probability of customer satisfaction by 0.915 times. Thus, it can be seen that the variable that influences customer satisfaction is the variable weight of goods which is based on partial testing and the greatest odds ratio.

Conclusion

Based on the analysis that has been carried out regarding customer satisfaction for loading and unloading containers at Tanjung Perak Port in Surabaya, it can be seen that the majority of the community, namely 86.6%, are satisfied with the existing services and the rest are still dissatisfied. Therefore, an analysis of the factors that influence satisfaction is carried out in order to optimize customer satisfaction. In this study, three basic variables were used, namely price of goods (X1), weight of goods (X2) and travel time (X3). These three variables have no relationship with each other, so there is no indication of a strong correlation between the three and it is found that the variable weight of goods (X2) has a significant influence on customer satisfaction based on partial test results and has the greatest odds ratio with the logit model is formed $\hat{g}(x) = 1.416 + 0.584X_2$ and this variable can increase customer satisfaction by 1.792 times. Therefore, the port should increase or optimize container capacity so that the public can increase the quantity of goods to be sent with services at Tanjung Perak Port. In addition, because this study is limited to using only three determinant variables, it is hoped that in the next similar study it can add other possible variables.

Acknowledgements

Thank you for the support for this research, Department of Marine Transportation Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Directorate General of Higher Education, Research, and Technology, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia.

References

[1] Saksono, H., Ekonomi biru: Solusi pembangunan daerah berciri kepu-lauan (studi kasus: kabupaten kepulauan anambas), Kementrian Dalam Negeri, Jakarta, 2013.

[2] Triatmodjo, B., Perencanaan pelabuhan (Edisi Pertama), Beta Offset, Yogyakarta, 2010.

[3] Pelabuhan Indonesia (Pelindo), Sejarah terbentuknya Pelindo, 2023. Available from: https://pelindo.co.id/page/sejarah-pelindo

[4] Badan Pusat Statistik (BPS), Volume dan nilai import menurut pelabuhan bongkar di provinsi jawa timur tahun 2020 dan 2021. Available from: https://jatim.bps.go.id/statictable/2022/06/03/2318/volume-dan-nilai-impor-menurut-pelabuhan-bongkar-di-provinsi-jawa-timur-2020-dan-2021.html

[5] Rizki, F., Widodo, D.A., Wulandari, S.P., Faktor-faktor risiko anemia gizi besi pada ibu hamil di jawa timur menggunakan analisis regresi logistic, *Jurnal Sains dan Seni ITS*. **4**(2) (2015) 305-310.

[6] Shafer, D.S., Zhang, Z., Introductory statistics, LibreTexts, p.11.1.2. Available from: https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Introductory_Statistics_(Shafer_and_Zhang)/11:_Chi-Square_Tests_and_F-Tests/11.01:_Chi-Square_Tests_for_Independence 2023.

[7] Starbuck, C., Logistic regression. In: Starbuck, C. The Fundamentals of People Analytics, *Springer, Cham*. (2023) pp. 1-371. DOI: https://doi.org/10.1007/978-3-031-28674-2_12

[8] Rusliyadi, M., Ardi, Y.W.Y., Winarno, K., Binary logistics regression model to analyze factors influencing technology adoption process vegetable farmers case in central java Indonesia. In: Abdullah, A.G. et al. (Eds.) Proceedings of the International Symposium Southeast Asia Vegetable 2021 (SEAVEG 2021), Yogyakarta, November 2021. Advanc-es in Biological Sciences Research, pp. 460-470. Atlantis Press, Amsterdam (2023). DOI: https://doi.org/10.2991/978-94-6463-028-2_48

[9] Johnson, N., Wichern, D., Applied multivariate statistical analysis sixth edition, Pearson Education Inc, New Jersey, 2007.

[10] Muzakkir, L., Analisis kinerja pelayanan pelabuhan penumpang tanjung perak, Undergraduate Thesis, Universitas Muhammadiyah Malang, 2020.

[11] Gujarati, D.N., Basic econometrics third edition, McGraw-Hill, Inc, New York, 2004.

[12] Guffey, D., Hosmer-Lemeshow goodness of fit test: Translations to the cox proportional hazards model. Thesis, University of Washington, 2012.

[13] Hosmer, D.W., Lemeshow, S., A goodness of fit test for the multiple logistic regression model, *Communication in Statistics.* **9**(10) (1980) 1043-1069.