International Journal of Offshore and Coastal Engineering
Vol.8 | No. 1 | pp. 17 – 22 | May 2024
e-ISSN: 2580-0914

Submitted: December 23, 2023 | Revised: March 3, 2024 | Accepted: April 12, 2024

Evaluation of International Shipping Alliance: A Case Study of Transport Capacity and Cost Efficiency Influence the Quality of Cooperation Inter-Country Using Discriminant Analysis

Tia Natalia[a*], Fadel Muhammad[a] and Tri Achmadi[a]

[a)] Institut Teknologi Sepuluh Nopember, Surabaya, Jawa Timur 60111, Indonesia
*Corresponding author: nata542@gmail.com

ABSTRACT

Shipping is a source of income that entrepreneurs increasingly demand because of the public's high use of this service, both for traveling and sending and transporting goods. Shipping companies offer their services domestically and cooperate with foreign shipping companies to increase revenue. In addition to increasing revenue, this business can increase the quality of cooperation between countries, which can be seen from transportation capacity and cost efficiency because more fleets will be provided, and infrastructure will be adequate because it consists of more than one country. Even though the shipping system cooperates with other countries and has an international standard, this does not necessarily satisfy the public with the service in the form of transportation capacity and fees charged. Not infrequently, many people complain that the costs incurred are quite high, but the implemented system is not following what most people expect. To see the quality of cooperation between countries (Y), this research will look at the transportation capacity factor (X1) and also cost efficiency (X2) using the discriminant analysis method. Discriminant analysis is a statistical method used to categorize an object into two or more groups based on the independent variable (X) used. The results obtained in this study show that the majority of people are satisfied with the existence of an international shipping alliance that is equal to 78%, with only cost efficiency (X2) being the main factor affecting customer satisfaction with the resulting classification accuracy rate of 71% and good models by 61%.

Keywords: *Cost Efficiency, Discriminant Analysis, International Shipping, Transport Capacity*.

1. INTRODUCTION

The world of shipping is a source of state revenue that is growing rapidly and has always been one of the modes of transportation used by most people to support their activities, both for going in and out of an area and trading goods on a big scale using containers. Containers are one of the choices that are most in demand by companies and individuals who want to send their production goods in large quantities because they have several advantages compared to ordinary packaging, including that containers can reduce the risk of goods sent being damaged because the material for making containers is thick metal. Hence, they are sturdy and safe. In addition, the loading and unloading process is much faster because it can be transported and moved using cranes and other tools. Another thing to consider is the ease of supervision because containers can be tagged in numbers, so they will not be lost or exchanged with containers owned by other companies [1]. Delivery of goods via containers is carried out between provinces and regions and can reach overseas due to the rapid development of the economic and trade system. Therefore, many countries take advantage of this situation by forming shipping alliances to facilitate the delivery of these goods via containers. An alliance is a bond between two or more countries with the political goal of mutual benefit or achieving common goals [2]. The shipping alliance that was formed was also intended to be able to overcome obstacles in exporting goods abroad. Indonesia is one of the countries involved in forming alliances with other countries, namely by forming the Indonesian Shipping Enterprises Alliance (Indonesian SEA) program, which aims to transport cargo for foreign markets [3]. Establishing Indonesian SEA follows President Joko Widodo's vision and mission through Presidential Instruction Number 5 of 2020 concerning Structuring the National Logistics Ecosystem, which aims to improve national logistics performance, improve the investment climate and increase the competitiveness of the national economy [3].

Even though the government has developed a shipping alliance system and added some infrastructure and shipping system development at several ports that support this system, it is possible that it will affect transport capacity and also cost efficiency incurred by companies and individuals who ship goods abroad, and This can affect customer or company satisfaction. Therefore, in this study, an analysis will be carried out using the discriminant method to determine customer satisfaction with the quality of interstate cooperation (Y), which will be seen from the transportation capacity factor (X1) and also cost efficiency (X2). Discriminant Factor Analysis (DFA), or simply Discriminant Analysis, is a statistical technique that aims to describe, explain and predict the membership of predefined groups (classes, modalities of the variable to be predicted...) of a set of observations (individuals, examples...) from a series of predictor variables like descriptors, exogenous variables, etc. [4].

2. METHODOLOGY

2.1 Discriminant Analysis

LDA is an extension of Fisher's discriminant analysis (FDA, Fisher, 1936), a multivariate method for finding a linear combination of continuous attributes best separating two classes. While FDA is a descriptive method used to assess the discriminative ability of the variables, LDA is used for class prediction [5]. There are generally two main approaches to Discriminant Analysis (depending on the objectives), where this method can be descriptive and predictive [4]. As a descriptive technique, the objective is to propose a new system of representation, latent variables formed from linear combinations of the predictor variables, which make it possible to discern groups of individuals as far as possible [4] and as a predictive, this technique is similar to supervised machine learning techniques such as decision trees, neural networks, etc. It is based on a probabilistic framework. The best known is certainly the multinomial distribution hypothesis (normal distribution). Added to the assumption of homoscedasticity, the conditional scatterplots have the same shape, and the author ends up with linear discriminant analysis [4]. The discriminant analysis model is an equation that shows a linear combination of various independent variables. With the equation,

$$Z = b_0 + b_1X_1 + b_2X_2 + \cdots + b_mX_m \qquad (1)$$

with,

Z = discriminant score
b = coefficient or discriminant weight
X= independent predictor or variable

2.2 Discriminant Assumption Testing

The discriminant assumptions used in this study are as follows.

1. Multivariate Normal Assumptions

A predictor variable can be said to have a normal multivariate distribution by visually looking at the Q-Q plot. If the plot's results describe a straight line, then the data can be declared normally distributed [6]. Mathematically, it can be calculated by referring to the previous hypothesis as follows.

H_0 : Data normally distributed multivariat e
H1 : The data is not normally distributed in multivariate

$$r_Q = \frac{\sum_{i=1}^{n}\left(x_{(i)} - \bar{x}\right)\left(q_{(i)} - \bar{q}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_{(i)} - \bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(q_{(i)} - \bar{q}\right)^2}} \qquad (2)$$

If the value of $r_Q < r_{table}$, then the decision is taken to reject H_0.

2. Homogeneity assumption

The homogeneity test is used to determine whether there are similarities between the several population variants used. The variance homogeneity test is necessary so that the differences are not caused by differences in the basis of the data used [7]. The hypothesis used is,

$$H_0 : \Sigma_1 = \Sigma_2$$
$$H_1 : \Sigma_1 \neq \Sigma_2$$

One test often used to test the assumption of multicollinearity is the Levene test, but this method often has deviations from the median [8]. Therefore, in this study, Box's M test will be used.

$$\text{Box's M} = MC^{-1} \qquad (3)$$

H_0 is rejected if, $p - value < \alpha$ or $\text{Box's M} > \chi^2_{\alpha,v}$.

3. Assumption of the Similarity of the Average Value Vector

Testing the assumption of similarity vectors needs to be done to see the similarity of the average vector values of all variables used with the hypothesis,

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

The average vector value equation calculation using the Barlett test is as follows [6].

$$V = \left(n - 1 - \frac{(p+g)}{2} \right) \ln \wedge^{*} \quad (4)$$

where,

n = Lots of observations
p = many predictor variables
k = many groups
$\wedge^{*}$ = Wilk's Lambda

H_0 is rejected $p - value < \alpha$ $V - Barlett > \chi^2_{\alpha, p(G-1)}$ for.

3. Multicollinearity Assumptions

Multicollinearity introduces instability in parameter estimation, resulting in imprecise estimates of factor loadings [9]. High correlations among predictors can cause parameter estimates to vary greatly across different model specifications or estimation techniques [10]. Several ways to detect multicollinearity are [11],

a. the value of the variance inflation factor (VIF) is more than 10
b. If there is a high correlation between the independent variables that is greater than 0.90
c. When the coefficient of determination R^2 is high, but there are no significant independent variables

2.3 Classification Accuracy

We can use the classification accuracy calculation for linear discriminant analysis based on the Apparent Error Rate (APER)

$$\text{Classification Accuracy} = (1 - APER) \times 100\%$$
$$= \left(1 - \left(\frac{n_{01} + n_{10}}{n} \right) \right) \times 100\% \quad (5)$$

to determine whether our classification results are correct. Table 1 explains classification information accuracy.

Table 1: Classification Accuracy

Actual	Prediction		Total
	0	1	
0	n_{00}	n_{01}	$n_{0.}$
1	n_{10}	n_{02}	$n_{1.}$
Total	$n_{.0}$	$n_{.1}$	n

2.4 Data Source

This study uses primary data from a survey of 50 people who have carried out international shipping, export and import activities. The variables used are the transportation capacity factor (X_1) and cost efficiency (X_2).

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

Descriptive statistics usually use a measure of data concentration and a measure of data distribution and present data in the form of tables, charts and graphs [12]. Table 1 shows the descriptive statistics for each predictor variable used in this study.

Table 1 provides information that the variance of the cost efficiency variable (X2) is very large because the range of the data is heterogeneous. Next is descriptive statistics for the response variable, namely customer satisfaction. Because it is categorical, this research will be presented as a bar chart in Figure 1.

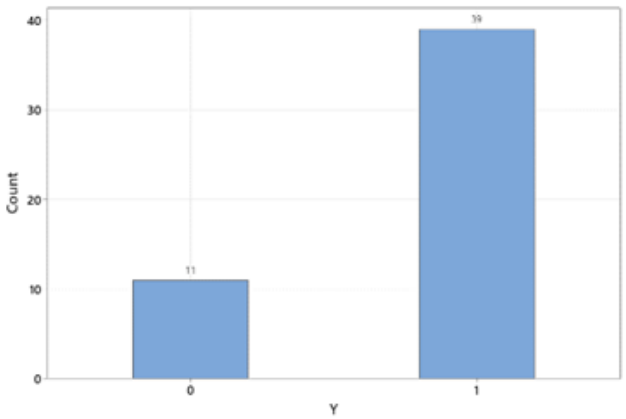

Figure 1: Response Variable Bar Chart

Figure 1 shows that the international shipping alliance has as many as 39 satisfied customers (category 1), or 78.0%. In comparison, the remaining 11 people, or equal 22.0%, are still dissatisfied with the existing capacity and cost-efficiency system.

3.2 Assumption Testing

Several assumptions must be met before carrying out an analysis using the discriminant method.

1. Normal Multivariate Assumption

Multivariate normal assumptions can be checked visually or by mathematical calculations using the hypothesis.

H0 : Data normally distributed multivariate
H1: The data is not normally distributed in multivariate

Visually, it can be seen based on the resulting Q-Q plot between the Mahalanobis distance and between the clusters, as shown in Figure 2.

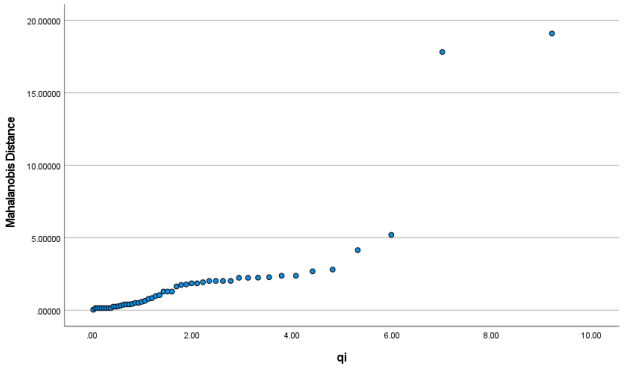

Figure 2: Q-Q plot

Figure 2 shows that the data is in the form of a straight line and has values that are quite close together in one regression line, so it can be said that the data meets the assumption of a multivariate normal distribution. Mathematically, the plot above can also be proven by calculating the correlation value between the distance to the cluster's center (Mahalanobis Distance) and between clusters (qi), as shown in Table 2.

Table 2: Test Results with Correlation

	Mahalanobis Distance	qi
Mahalanobis Distance	1	**0.846**
qi	**0.846**	1

Based on Table 2, it is known that the correlation coefficient between the cluster center distance (Mahalanobis Distance) and the distance between clusters (qi) is 0.846, where this value is greater than 0.8, which indicates a very high correlation or a high relationship so that the data can be said to be distributed multivariate normal.

2. Homogeneity Assumption

Analysis with a good discriminant is when the data used tends to be homogeneous, or the variance of the data is the same or not diverse. Testing the assumption of homogeneity in this study was carried out to prove that the data used has the same variance as the hypothesis used in this study,

$$H_0 : \sum_1 = \sum_2$$
$$H_1 : \sum_1 \neq \sum_2$$

This study used Box's M test to test the assumption of homogeneity, and the resulting p-value was 0.267. By comparing this value with a significance level of 5%, it was found that the p-value > 0.05, which means it failed to reject H0 so that the conclusion of variance from the data used is the same and has fulfilled the assumptions of homogeneity.

3. Assumption of Similarity Vector Mean Value

The assumption of the similarity of the average value vector needs to be met before carrying out an analysis using the discriminant to find out whether the average value vector already has similarities with the hypothesis used,

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

The test results using Wilks' lambda were obtained as shown in Table 3.

Table 3: Wilks' Lambda Test Results

Wilks' Lambda	P-value
0.939	**0.227**

Tests carried out to test the similarity of the average value vector using the Wilks' Lambda test are presented in Table 3, which shows that the resulting p-value is 0.227, where this value is greater than 0.05, which gives the conclusion that it fails to reject H0, which means that the vector is the average value of the data. The research is the same, so it has met the assumption of similarity with the average value vector.

4. Multicollinearity Assumptions

The last assumption that must be met is the assumption of multicollinearity, where the data used should not have a relationship with each other. The VIF value can detect the multicollinearity assumption. In this study, a VIF value of 1.157 was obtained for each variable, where data is said to have no relationship with each other if the VIF value is less than 10. Thus, this study's multicollinearity assumption has been fulfilled because the resulting VIF <10.

3.3 Discriminant Analysis

Discriminant analysis is very useful where the total sample can be divided into groups based on the characteristics of known variables and is also used to see differences between groups where the response variable is categorical or qualitative data and more than one predictor variable is in the form of quantitative data. In this study, the Discriminant Analysis Test Results can be seen in Table 4, in which the discriminant function formed is $Z=-0.982-0.175X^1$, where this function shows that only the transport capacity variable (X1) is included in the model.

Table 4: Discriminant Analysis Test Results

Variable	Canonical Discriminant Function
X_1	-0.175
X_2	0.000
Constant	-0.982

3.4 Model Goodness

The goodness of the model is useful for knowing how well the model can explain all the variables in the study. In this study, the model's goodness was measured by the canonical correlation obtained when carrying out discriminant analysis, as shown in Table 5.

Table 5: Canonical Correlation

Function	Canonical Correlation
1	0.247

The canonical correlation formed in this study is 0.247, which means that the formed discriminant function can explain the variability of the categories of origin of the region by 61%. In comparison, the remaining 39% is explained by other variables not used in the process analysis. Furthermore, it can be seen that the contribution of each variable in the discriminant function that is formed by looking at the loading discriminant value was obtained, as shown in Table 6.

Table 6: Discriminant Loading

Function	Discriminant Loading
X_1	-0.087
X_2	0.887

Variables with an absolute discriminant loading value > 0.40 contribute significantly to classification. The cost efficiency variable (X2) determines customer satisfaction in this analysis.

3.5 Model Classification

The variables used in the study must be classified into the response variable group, namely the satisfied category, denoted by one, and the dissatisfied category, denoted by 0 (Table 7).

Table 7: Function at group centroid

Category of Y	Function
Dissatisfied	0.471
Satisfied	-0.133

The prior probability assumptions are assumed to be identical, and the following classification calculations are obtained.

$$Z = \frac{Z_A + Z_B}{2} = \frac{0.471 + (-0.133)}{2} = 0.169$$

Based on the calculation results, it is known that if the discriminant score obtained from the discriminant function is <0.169, then it will be classified in category 0, namely dissatisfaction. In contrast, if the discriminant score obtained from the discriminant function is > 0.169, it will be classified in category 1, namely the satisfied category. The classification results obtained are shown in Table 8.

Table 8: Classification Result

Category of Y	Predicted Group Membership		Total
	Dissatisfied	Satisfied	
Dissatisfied	8	3	11
Satisfied	21	18	39

Table 8 shows that the number of people who are dissatisfied and appropriately grouped in the dissatisfied category is as many as eight people or 72.7%, while those who are dissatisfied but categorized as satisfied are as many as three people or equal to 27.3%. On the other hand, the number of people who were satisfied with the service and correctly categorized into the satisfied group was 18 people or equal to 46.2%, while people who were initially satisfied but were categorized into the dissatisfied category were as many as 21 people or equal to 53.8% so, to calculate the APPER value is,

$$APER = \frac{21+3}{8+3+21-18} = \frac{24}{14} = 1.71$$

$$Accuration = 1 - APER = 1 - 1.71 = |-0.71| = 0.71$$

Based on the discriminant function formed, it is known that the resulting accuracy value is 71%, where this accuracy value is large enough so that the discriminant function can be said to be good enough in classifying customer satisfaction.

4. CONCLUSIONS

Based on the previous analysis, it is known that the majority of the public is satisfied with the existence of an international shipping alliance, namely 78%, which is related to cargo capacity and also costs to be incurred where these two variables can be differentiating variables that differentiate customer satisfaction because these variables are known has a value of F < 0.05. However, only cost is the main factor that influences customer satisfaction. The resulting classification accuracy rate is 71%, and the model's goodness is 61%. Because the variable price of goods is a factor that influences customer satisfaction, this can be used as an evaluation material for shipping service companies and the government to be able to reduce prices or costs incurred by customers to send goods abroad.

Future research is expected to develop other methods that can explain customer satisfaction using transport capacity variables (X_1) and cost efficiency (X_2) because the discriminant method in this study, one data point, still cannot be grouped. In addition, it is hoped that further research will be able to add other variables so that it is known that variables other than the transport capacity factor (X_1) affect customer satisfaction related to the existence of international shipping alliances.

ACKNOWLEDGEMENTS

The authors thanks the 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 for supporting this research.

REFERENCES

1. Patiung, D., Brasit: Peranan petikemas dalam menunjang ekspor impor komoditi hasil pertanian di pelabuhan soekarno hatta makassar. Universitas Hasanudin, Makassar (2011)
2. Kamus Besar Bahasa Indonesia. https://kbbi.web.id/aliansi (2023)
3. Thenniarti, D.: Mengatasi hambatan ekspor di pelayaran kemhub bentuk Indonesian sea dan seacomm. Info Publik. https://www.infopublik.id/kategori/nasional-ekonomi-bisnis/ (2023).
4. Ontshick, L. L., Sabue, J. C. M., Kiangebeni, P. M., Mbaya, O. T., Ntamabyaliro, J. M. N., Tamfumu, J. J. M., Matendo, R. M. M.: Comparison of the performance of linear discriminant analysis and binary logistic regression applied to risk factors for mortality in ebola virus disease patients. Journal of Electronics, Electromedical Engineering, and Medical Informatics. 5 (3), 134-142 (2023). https://doi.org/10.35882/jeemi.v5i3.303
5. Graf, R., Zeldovich, M., Friedrich, S.: Comparing linear discriminant analysis and supervised learning algorithms for binary classification-A method comparison study. Biometrical Journal, pp. 1-20 (2022). https://doi.org/10.1002/bimj.202200098
6. Rusli, Gaffar, S. B., Jasruddin, Ahmad, M. A.: Pengujian normal multivariat dan homoskedastisitas matriks varians-kovarians pada prestasi balajar dan kredit lulus mahasiswa unm angkatan 2014, pp. 2. Universitas Negeri Makassar, Makassar (2018)
7. Usmadi: Pengujian persyaratan analisis (uji homogenitas dan uji normalitas). Jurnal Inovasi Pendidikan. 7 (1), 51 (2020)
8. Brown, M. B., Forsythe, A. B.: Robust tests for the equality of variances. Journal of the american statistical association. 69, 364-367 (1974)
9. Kyriazos, T., Poga, M.: Dealing with multicollinearity in factor analysis: the problem, detections and solutions. Open Journal Statistics. 13 (3), 404-424 (2023). https://doi.org/10.4236/ojs.2023.133020
10. Mindrila, D.: Maximum likelihood and diagonally weighted least square estimation procedures: a comparison of estimation bias with ordinal and multivariate non-normal data. International Journal of Digital Society. 1, 60-66 (2010). https://doi.org/10.20533/ijds.2040.2570.2010.0010
11. Ghozali, I.: Aplikasi analisis multivariate dengan program spss. Badan Penerbit UNDIP, Semarang (2017)
12. Walpole, R. E.: Pengantar statistika. PT. Gramedia Pustaka Utama, Jakarta (1995)