JOMTA
Journal of Mathematics: Theory and Applications
Vol. 7, No. 2, 2025, pp. 97 - 106
P-ISSN: 2685-9653 e-ISSN: 2722-2705
DOI: 10.31605/jomta.v7i2.5359

Classification of Regencies/Cities in South Sulawesi Province Based on Business Sectors Using Discriminant Analysis

Isma Muthahharah[1*], Farida Islamiah[2]

[1] Statistics,Universitas Negeri Makassar, 90224, Indonesia
[2] Busniss Digital, Universitas Negeri Makassar, Makassar, 90221, Indonesia

Corresponding email: isma.muthahharah@unm.ac.id*

Abstract

This study employs the Linear Discriminant Analysis (LDA) method to classify districts and cities in South Sulawesi Province according to their dominant economic sectors. Data obtained from the Central Statistics Agency (BPS) for the period 2019–2023 were analyzed, with the Gross Regional Domestic Product (GRDP) quartiles serving as the dependent variable, and the Labor Force Participation Rate (LFPR), Number of Business Entities (NBE), and Open Unemployment Rate (OUR) as independent variables. The results indicate that LFPR and OUR are the most influential factors in differentiating regional groups. Although the model met the assumptions of normality, homogeneity of covariance, and absence of multicollinearity, its classification accuracy was relatively low (37.5%). This finding suggests that model improvement is needed, either by incorporating additional explanatory variables or by employing more advanced classification approaches. The study provides meaningful insights to support regional economic development and policy formulation in South Sulawesi Province.

Keywords: *classification; gross regional domestic product (GRDP); labor force participation rate (LFPR); linear discriminant analysis (LDA); open unemployment rate (OUR)*

Received 07-08-2025 Revised :07-10-2025 Accepted :14-10-2025 Published :17-10-2025

1. Introduction

Social transformation can be reflected in the equitable distribution of socio-economic resources such as education, health, housing, clean water, and political participation, while cultural transformation is often associated with the rise of nationalism and changes in societal values and norms, such as a shift from traditional to more modern and rational institutions [1]. These transformations are closely linked to economic development and structural changes across sectors. The Gross Regional Domestic Product (GRDP), which represents the total value added generated by all production activities within a region, serves as an important indicator of regional economic performance. GRDP at constant prices is commonly used to measure economic growth over time [2].and understanding variations in GRDP across regions helps identify patterns of regional development and disparities. In this context, this study aims to classify districts and cities in South Sulawesi Province according to their dominant economic sectors using the Linear Discriminant Analysis (LDA) method. The analysis seeks to determine the socio-economic factors that most significantly differentiate regional economic performance, thereby supporting more targeted development planning and policy formulation.

The role of GRDP by business sector in Indonesia reflects the contribution of each economic sector to the country's economy. Referring to the business sector, GRDP displays the contributions of various economic sectors in Indonesia. Each of these sectors plays a different role in the Indonesian economy, and its contribution to GRDP reflects the country's economic dynamics [3]. A country's high Gross Domestic Product (GDP) does not always indicate equitable income distribution. Facts show that people's income is often not distributed fairly, and there is even a tendency to show the opposite. Inequality in income distribution can lead to disparities[3].

This study employs the discriminant analysis method, a statistical technique used to classify regions into distinct groups based on several influencing variables. In the context of regional economic analysis, the

Gross Regional Domestic Product (GRDP) and its sectoral components reflect the structure and growth dynamics of local economies. Given the complexity and interdependence among these sectors, discriminant analysis is particularly appropriate for identifying and distinguishing regions with similar economic characteristics [4]. In the context of object classification, this method functions to assign each observation into one of two or more predefined classes. In discriminant analysis, each object can belong to only one group, making the classification mutually exclusive. After the grouping process, discriminant analysis is also used to evaluate the accuracy of the classification performed. The method used in this study is Discriminant Analysis to classify Regencies/Cities in South Sulawesi Province based on Business Field Indicators from 2019 to 2023 [5].

Several previous studies have also explored comparisons of this method. A study by [6] examined the use of Discriminant Analysis with K-Fold Cross Validation for Water Quality Classification in Pontianak. Meanwhile, research conducted by [7] used Discriminant Analysis to classify Regencies/Cities in South Sulawesi Province based on Human Development Index (HDI) indicators. Based on these two studies, the authors were motivated to conduct research on the classification of Regencies/Cities in South Sulawesi Province according to business fields using Discriminant Analysis. Unlike previous studies that applied Discriminant Analysis to water quality and the Human Development Index, this study applies the method to Gross Regional Domestic Product (GRDP) data by business sectors, offering a new contribution to understanding the regional economic classification and structural differences among regencies and cities in South Sulawesi.

2. Theoretical Reviews

2.1 Discriminant Analysis

Discriminant analysis is a statistical technique used to classify items or observations based on distinguishing characteristics. Although this technique is comparable to multiple linear regression (also known as multivariate regression), its application differs because discriminant analysis is used when the independent variables are numerical data (interval or ratio scale), and the dependent variable is categorical (nominal or ordinal scale) [8]. Multiple regression, on the other hand, is used when the dependent variable may be either categorical or numerical, and the independent variables are numerical. Discriminant analysis refers to a statistical technique for classifying individuals into groups based on a set of random variables that are strongly and independently related to one another[9].

2.2 Discriminant Assumption

a. Multivariate normality requires that the independent variables follow a normal distribution. If the data are not normally distributed, the accuracy of the discriminant model may be affected. When the independent variables clearly do not follow a normal distribution, alternative methods such as logistic regression may be used[10].

$$W_x = \frac{\tilde{\sigma}_x^2}{S_X^2} \tag{1}$$

This assumption will be rejected at the significance level α distribution if the test statistics $Wx < k\alpha$ with $k\alpha$ distribution Wx. If the data do not follow a normal distribution, the Box-Cox test can be used to find a transformation that makes the data more closely approximate a normal distribution.

b. Homogeneity of Covariance Matrices. The covariance matrices of the independent variables must be equal or homogeneous across groups. If this assumption is violated, the results of the analysis may become unreliable[11].

c. No Multicollinearity: Independent variables should not be highly correlated with one another. If two independent variables exhibit a strong correlation, multicollinearity occurs. This can distort model interpretation, similar to what happens in multiple regression analysis[12].

d. No Extreme Outliers: There should be no extreme outliers in the independent variables. If outliers are included in the analysis, they may reduce the accuracy of the model in performing classification [13].

2.3 Discriminant Model

The general form of the discriminant model is as follows [14]:

$$D_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_p x_{ip} \quad (2)$$

Where D_i is the discriminant score for the i observation, b_p is the coefficient of the discriminant function, and x_{ip} is the value of the p variable for the i observation.

2.4 Gross Regional Domestic Product (GRDP)

One of the main indicators of a region's economic health over a specific period, both at current and constant prices, is the Gross Regional Domestic Product (GRDP). GRDP includes several sectors in its measurement, one of which is the economic sector or field of business. The field of business refers to economic activities carried out by companies or individuals to produce goods and services for the purpose of generating profit. GRDP by economic sector plays a crucial role in determining the economic growth rate of a region [15]. Analyses using GRDP data by economic sector are widely conducted to assess the growth and economic disparities among regions. Another important metric for evaluating the effectiveness of regional development is economic growth. Gross Regional Domestic Product (GRDP), which represents the total value of goods and services produced in a region over a certain period, is frequently used for this purpose. In addition to aiding regional development policy-making, GRDP provides data on the relative contribution of each economic sector to regional economic growth [16].

Table 1. GRDP Criteria by Economic Sector

Quartile	GDRP Range	Category
1	GDRP < 25%	Low
2	25% < GDRP ≤ 50%	Lower _ Middle
3	50% < GDRP ≤ 75%	Upper _ Middle
4	GDRP > 75%	High

(Source: Badan Pusat Statistik)

Based Table 1 presents the classification criteria for the contribution levels of business sectors to the total GRDP in each regency/city of South Sulawesi. The "Low" category indicates sectors contributing less than 25% to total GRDP, while the "Lower Middle" and "Upper Middle" categories represent sectors with moderate economic roles contributing between 25% and 75%. The "High" category identifies dominant sectors contributing more than 75% to total GRDP. This quartile-based classification serves as the basis for discriminant analysis, which aims to group regencies and cities according to the contribution patterns of their main economic sectors.

3. Methods

3.1. Types of Research

This study is a quantitative research that employs both descriptive and inferential approaches. Quantitative research aims to answer research questions by applying a structured plan and adhering to the principles of scientific inquiry. In this context, descriptive analysis is used to describe or explain the collected data without attempting to generalize the findings. In contrast, inferential analysis is a statistical technique that examines data from a sample and allows for the extrapolation of findings to a larger population. By combining these two approaches, the research provides a more in-depth understanding of the studied phenomena.

3.2 Data Source

The primary data source in this study is the Statistics Indonesia (BPS) of South Sulawesi Province, covering 24 regencies and cities from 2019 to 2023. GRDP data by business field were obtained and

expressed as the percentage contribution of each sector to the total GRDP, then averaged over the five-year period to represent stable economic characteristics for each region. The study uses the classification of regencies/cities based on dominant economic sectors as the dependent variable (Y), while the independent variables include GRDP by economic sector, Labor Force Participation Rate (LFPR), Number of Business Enterprises (NBE), and Open Unemployment Rate (OUR).

3.4 Data Analysis Techniques

This study employs the Linear Discriminant Analysis (LDA) technique to classify regions based on Economic Sector (Y), using three independent variables: Labor Force Participation Rate (LBRF) (X_1), Number of Business Enterprises (NBE) (X_2), and Open Unemployment Rate (OUR) (X_3). LDA was selected for its ability to distinguish groups with similar economic patterns and validate classification accuracy. The analysis was performed using R software to ensure the consistency and robustness of the results. The steps involved in the data analysis technique used in this study are as follows:

a. Conducting univariate normality tests using the Shapiro-Wilk test to ensure that each predictor variable follows a normal distribution.

b. Applying Box-Cox transformation if the data do not follow a normal distribution.

c. Re-testing normality after data transformation to confirm whether the data distribution has improved.

d. Performing homogeneity test of covariance matrices using Box's M test to check whether the group covariances are homogeneous.

e. Conducting a multicollinearity test using Variance Inflation Factor (VIF) to ensure no high correlation exists among the predictor variables.

f. Performing discriminant analysis by building a discriminant model to classify regencies/municipalities based on economic sector.

g. Conducting prediction and model evaluation, using the constructed model to predict the categorical classification of regencies/municipalities and assess its accuracy.

h. Visualizing the results of LDA by plotting function scores to observe group separation based on the analysis results.

4. Results and Discussion

4.1 Testing the Assumptions of Discriminant Analysis

4.1.1 Univariate Normality Test

The univariate normality test is used to determine whether a single variable is normally distributed.

a. Shapiro-Wilk Test

The Shapiro-Wilk method is one of the alternative procedures used to test the normality of a dataset. It was developed based on the expected values of a standard normal distribution and the sample means that have been applie

Table 2. Results of the Univariate Normality Test

LBRF	NBE	OUR
0,7099	0,000004	0,00342

The results of the Shapiro-Wilk normality test show that the LFPR variable is normally distributed (p =

0,7099), whereas NBE (p = 0,000004) and OUR (p = 0,00342) are not normally distributed, as their p-values are less than 0.05. This indicates that the normality assumption is satisfied for LFPR, but not for NBE and OUR. If further analysis requires normally distributed data, data transformation or the use of non-parametric methods can be considered as alternatives.

b. Box-Cox Test

The Box-Cox transformation is a type of power transformation applied to the response variable, developed by Box and Cox. Its main purposes are to normalize the data, linearize the regression model, and stabilize the variance.

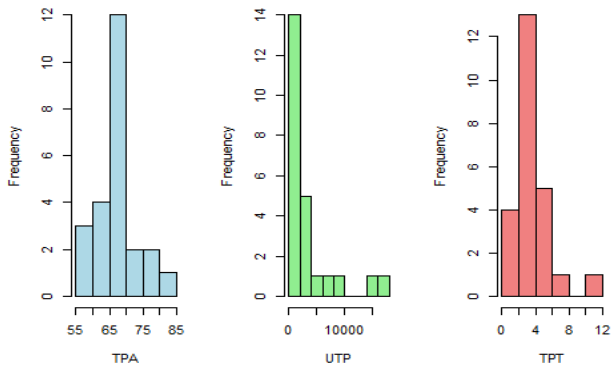

Figure 1. Box-Cox transformation

The Box-Cox transformation produced a symmetrical and nearly normal distribution for the LFPR histogram, indicating the success of the transformation process. The NBE histogram showed consistent variation and also appeared to approach a normal distribution, suggesting its potential for further analysis. On the other hand, the OUR histogram showed significant improvement and displayed characteristics aligned with a normal distribution. Overall, all variables demonstrated progress toward meeting the normality assumption, enabling the possibility for more in-depth analysis.

4.1.2 Homogeneity Test of Covariance Matrices

a. Box's M Test

Box's M test was applied to assess the equality of covariance matrices among groups. A non-significant result ($p > 0.05$) indicates homogeneity of covariance matrices, fulfilling the assumption for discriminant analysis, while a significant result ($p < 0.05$) denotes heterogeneity among groups.

Table 3. Box's M Test

Chi-Square	Df	p-value
23,139	18	0,1853

The results of Box's M test show a Chi-Square value of 23,139 with degrees of freedom (df) = 18 and a p-value of 0,1853. This test is used to evaluate the homogeneity of covariance matrices across groups, where the null hypothesis (H_0) states that the covariance matrices are homogeneous. Since the p-value is greater than 0,05, there is no significant difference in covariance matrices among the groups, indicating that the assumption of homogeneity is satisfied. Therefore, further analyses that rely on this assumption, such as discriminant analysis or MANOVA, can be conducted without concerns regarding group variability.

b. Multicollinearity Test

The multicollinearity test is a method used to identify linear relationships among independent variables in multiple regression analysis. The purpose of this test is to evaluate the degree of correlation between

each pair of independent variables. A good regression model should exhibit no significant correlation among the independent variables.

Table 4. Multicollinearity Test

LFPR	NBE	OUR
1.178375	1.217327	1.226786

Based on the analysis, the Variance Inflation Factor (VIF) values for each variable are as follows: LFPR = 1.178, NBE = 1.217, and OUR = 1.227. Since all VIF values are below 5, this indicates the absence of significant multicollinearity among the predictor variables. Therefore, these variables can be used in subsequent discriminant analysis without the risk of distortion due to high intercorrelations.

4.2 Discriminant Analysis

4.2.1 Prior Probability

Prior probability is a concept in statistics, particularly within the Bayesian approach, that reflects the degree of belief in the likelihood of an event or hypothesis being true before observing any new data.

Table 4. Prior Probabilities

1	2	3	4
0.25	0.25	0.25	0.25

The prior group probabilities in Linear Discriminant Analysis (LDA) indicate that each group has an equal initial probability of 0.25 or 25%. This means that, before considering the predictor variables (LFPR, NBE, and OUR), each observation is assumed to have an equal chance of belonging to one of the four groups.

4.2.2. Group Means

Group means refer to the average (mean) values of a variable within each distinct group. They serve to highlight the differences in characteristics among the groups. These differences are a critical foundation in discriminant analysis, as they reflect how well the variables can distinguish one group from another.

Table 5. Group Mean

LBRF	NBE	OUR
693,0833	1824,333	4,406667
68,66333	3139,333	3,008333
64,69167	3640,000	2,778333
65,28667	4121,833	5,170000

The group means in the discriminant analysis indicate differences in LFPR, NBE, and OUR across the groups. Group 1 exhibits the highest LFPR and the lowest NBE, while Group 4 shows the highest NBE and OUR. Group 3, in contrast, has the lowest LFPR and OUR, with its NBE value exceeding those of Groups 1 and 2. It should be noted that the values presented are the original means, not transformed by any Box-Cox or other normalization procedure. This variation in group means highlights patterns that can be exploited in the classification process, with each variable contributing to differentiating the groups based on their economic characteristics.

4.2.3 Linear Discriminant Coefficients

The Linear Discriminant Analysis (LDA) produced two discriminant functions used to classify regencies/cities based on their economic characteristics. The standardized discriminant function coefficients are presented Table 6.

Table 6. Linear Discriminant Coefficients

	LD1	LD2	LD3
LBRF	-104.8542	309.5939	112.8661
NBE	-0.1926	-0.4305	0.3726
OUR	-3.0723	-3.0723	0.3937

a. Linear discriminant coefficients are values used to construct the linear discriminant functions that form linear combinations of the predictor variables (such as LFPR, NBE, and OUR), aiming to maximize the separation between groups. These coefficients indicate the contribution of each variable in distinguishing among the groups.
b. The analysis results show that the Open Unemployment Rate (OUR) and Labor Force Participation Rate (LFPR) have the greatest influence, particularly on LD1 and LD2, which together explain most of the variability in the data. OUR has the highest weight on LD1, serving as the primary factor in the initial group separation, while LFPR contributes more dominantly to LD2 with the highest coefficient value (309.5939), indicating its significant role in differentiating groups along the second dimension. However, since the Conclusions section states that LFPR and NBE contribute most significantly, it is recommended that the authors re-verify the analysis results to ensure consistency throughout the manuscript.
c. The third discriminant function (LD3) displays a more balanced contribution, where LFPR remains the dominant variable, but NBE and OUR begin to show more notable positive influence. Nevertheless, overall, NBE has a smaller impact compared to LFPR and OUR, though it still contributes to the classification process.
d. Thus, it can be concluded that LFPR is the most influential variable overall, followed by OUR, while NBE provides additional complementary contributions. Together, these three variables form a discriminant model that can be used to classify the data optimally.

4.2.4 Proportion of Trace

The proportion of trace indicates the extent to which each discriminant function contributes to distinguishing between groups. This information is useful for identifying which functions are most significant in the classification process. A higher proportion value for a particular function implies a greater role in explaining group differences. By examining these proportions, researchers can determine the relative importance of each discriminant function in the overall model and focus on the most impactful dimensions for interpretation and classification.

Table 7. Proportion of Trace

LD1	LD2	LD3
0.5723	0.3895	0.0382

The analysis results indicate that LD1 contributes the most is 57,23%, making it the primary component in group separation. LD2 accounts for 38,95%, indicating a still significant role in group discrimination. Meanwhile, LD3 explains only 3,82 %, suggesting its influence in distinguishing between groups is minimal. Therefore, the first two components (LD1 and LD2) are sufficient to optimally explain the data classification pattern.

4.3 Prediction and Model Evaluation

4.3.1 Distribution of Discriminant Scores

The distribution of discriminant scores illustrates the spread of values obtained from the discriminant functions (such as LD1, LD2, etc.) for each individual in the dataset. These values are calculated by computing a linear combination of the predictor variables.

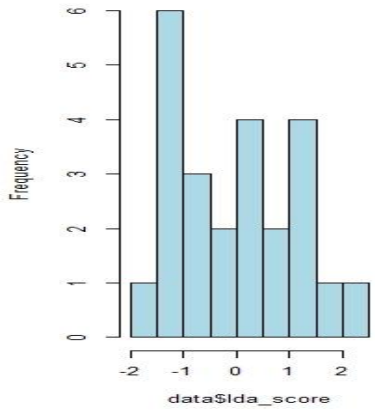

Figure 2. Distribution of Discriminant Scores

Based on the results of the Linear Discriminant Analysis (LDA) (LDA scores), the result show the scores are spread within the range of -2 to 2, with varying frequencies across intervals. This score distribution indicates that the discriminant model has successfully mapped the data based on the combination of the predictor variables used. The distribution appears relatively symmetric, with no extreme skewness, suggesting that the group separation is fairly balanced. Additionally, the highest frequency is observed around the score of -1, indicating that most of the samples have discriminant scores leaning towards negative values, although some positive values are still present.

This distribution confirms that the discriminant model is capable of distinguishing between groups based on the analyzed variable patterns. However, to ensure the model's effectiveness in optimally separating the groups, further evaluation is recommended.

4.3.2 Classification Table of Model Prediction Results

The classification results table is a method for evaluating how accurately a predictive model can identify or assign data into the correct categories. the numbers in each cell represent the number of regencies/cities predicted by the LDA model for each class. This clarification helps readers understand that the analytical unit in this study is the regency/city.

Table 8. Model Prediction Results

Category	Model Prediction				
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Akurasi(%)**
Class 1	2	1	1	2	33,3%
Class 2	2	0	2	2	0%
Class 3	1	0	4	1	66,7%
Class 4	2	0	1	3	50%

The LDA model used in this classification yielded an overall accuracy of 37.5%, indicating that the model is not yet effective in distinguishing between classes in the dataset. Class 3 showed relatively good classification accuracy at 66.7%, while Class 1 and Class 4 achieved 33.3% and 50% accuracy, respectively. However, Class 2 failed to be correctly classified at all, with an accuracy of 0%.

4. 4 Visualization of LDA Results

Visualization in Linear Discriminant Analysis (LDA) provides an intuitive understanding of how well the model separates groups based on discriminant functions. The most common visualization is a scatterplot of the first two discriminant scores (LD1 and LD2), where each point represents an observation in the discriminant space. The inclusion of LDA tables and graphs helps readers better interpret and understand the model's classification results.

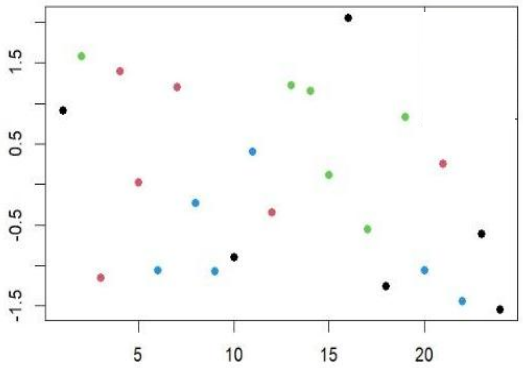

Figure 3. Visualization of LDA Results

Each class in the LDA scatterplot is represented using distinct colors: Class 1 in black, Class 2 in pink, Class 3 in green, and Class 4 in blue. Ideally, if the LDA model performs well, data points from the same class should cluster closely together with minimal overlap across groups. However, the visualization shows considerable overlap, particularly between Class 2 and other classes, suggesting that the model struggles to distinguish Class 2 observations effectively. This pattern is consistent with the low overall classification accuracy (37.5%), indicating that the discriminant functions may not fully capture the underlying differences among the groups. The poor separation of Class 2 may be due to similarities in predictor variables or overlapping characteristics with other classes, which limits the model's discriminating power.

4. Conclusions

The LDA model shows both strengths and limitations in classifying regional economic data. While assumptions of normality and covariance homogeneity are largely met and LBRF and NBE emerge as key discriminating variables, the model's overall accuracy remains low (37.5%), with Class 2 entirely misclassified. This limitation may result from a small sample size and limited predictor variables, which restrict the model's ability to capture complex regional dynamics. Despite this, the findings offer practical insights for policymakers by identifying economic indicators most relevant for regional differentiation. Future studies should expand the dataset and explore advanced classification methods—such as Random Forest or SVM—to enhance predictive performance and strengthen the model's policy relevance.

References

[1] Suradi, "Pertumbuhan Ekonomi dan Kesejahteraan Sosial (Economic Growth And Sosial Welfare)," *Puslitbang Kesejaht. Sos.*, vol. 17, no. 200, 2012.

[2] M. Jamilah, A. Hasibuan, A. Rusgiyono, and D. Safitri, "Pemodelan Produk Domestik Regional Bruto (PDRB) Di Provinsi Jawa Tengah Menggunakan Bootstrap Aggregating Multivariate Adaptive Regression Splines," vol. 8, pp. 139–149, 2019.

[3] R. Johanes and D. Takari, "A review of the tourist industry ' s role in Indonesia ' s GDP growth,"

vol. 10, no. 1, pp. 8–17, 2024.
[4] F. Hayati, D. Vionanda, Y. Kurniawati, and T. O. Mukhti, "Comparison of Linear Discriminant Analysis with Robust Linear Discriminant Analysis," vol. 2, pp. 353–359, 2024.
[5] S. H. Mahmood and H. H. Khider, "Comparison of Linear Discriminant Analysis and Artificial Neural Networks for Stroke patients Classification Stroke," vol. 13, no. 2, pp. 208–220, 2023.
[6] L. Mardiana, D. Kusnandar, and N. Satyahadewi, "Analisis Diskriminan Dengan K Fold Cross Validation Untuk Klasifikasi Air di Kota Pontianak," vol. 11, no. 1, pp. 97–102, 2022.
[7] I. Muthahharah, S. M. Meliyana, and A. S. Ahmar, "K Means Cluster for Grouping Regencies / Cities in South Sulawesi Province Based Human Development Index on the 2023," vol. 4, no. 3, pp. 357–364, 2024.
[8] D. Kertayuga, E. Santoso, and N. Hidayat, "Prediksi Nilai Ekspor Impor Migas dan Non-Migas Indonesia Menggunakan Extreme Learning Machine (ELM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 6, pp. 2792–2800, 2021.
[9] B. J. Babin and R. E. Anderson, *Multivariate Data Analysis*. 2014.
[10] P. G. Subhaktiyasa, "PLS-SEM for Multivariate Analysis : A Practical Guide to Educational Research using SmartPLS," vol. 4, no. 3, 2024.
[11] K. F. Nimon, "Statistical assumptions of substantive analyses across the general linear model : a mini-review," vol. 3, no. August, pp. 1–5, 2012.
[12] M. J. Naufal, D. P. Ompusunggu, R. A. Sinaga, M. Dola, A. Sitohang, and T. N. Gunawan, "A Theoretical Study of Multicollinearity and Linearity in Econometric Models for Economic Research," vol. 21, no. 1, pp. 43–52, 2025.
[13] C. Sanjeev, K. Dash, A. Kumar, S. Dehuri, and A. Ghosh, "An outliers detection and elimination framework in classification task of data mining," *Decis. Anal. J.*, vol. 6, no. May 2022, p. 100164, 2023, doi: 10.1016/j.dajour.2023.100164.
[14] M. O. Adebiyi, M. O. Arowolo, M. D. Mshelia, and O. O. Olugbara, "applied sciences A Linear Discriminant Analysis and Classification Model for Breast Cancer Diagnosis," 2022.
[15] M. Y. Majo, "Unveiling Human Development Index Mediation on Consumption Dynamics," vol. 13, no. 1, 2024.
[16] B. S. Utomo and J. Timur, "Analysis of the Quality of Economic Growth in Districts / Cities in East Java," vol. 3, no. 3, pp. 699–712, 2024.