Discriminant Analysis for Determination of Early Childhood Education Accreditation in Southeast Sulawesi Province

Makkulau[1],*, Gusti N.A. Wibawa[1], Andi Tenri Ampa[1], Andi Tenri Pannangngareng Makkulau[1], Sri Harini[2], Angga Dwi Mulyanto[2]

[1]Statistics Study Program, Universitas Halu Oleo, Indonesia
[2]Department of Mathematics, UIN Maulana Malik Ibrahim Malang, Indonesia

Email: kulau.statistika@gmail.com

ABSTRACT

Discriminant Analysis is a statistical analysis that can classify cases on independent variables into groups or categories of dependent variables. The main objective of this research is classify eight indicators of the National education standards (SNP) early childhood and classify the accreditation value of early childhood (PAUD) in Southeast Sulawesi Province. The method used in this study used discriminant analysis. Accreditation value factors used in this study include Standards for Child Development Achievement Levels (X_1), Content Standards (X_2), Process Standards (X_3), Standards for Educators and Education Personnel (X_4), Facilities and Infrastructure Standards (X_5), Management Standards (X_6), Financing Standards (X_7) and Education Assessment Standards (X_8). Based on the results of data analysis, 8 SNP Indicators qualify as a form of discriminant equation model and accreditation value obtained based on the calculations of the National Accreditation Organization (BAN) PAUD and non Formal Education (PNF) Southeast Sulawesi are classified as follows: there are divided into 3 classifications, namely Accreditation C is 91.7%, Accreditation B is 85.1%, and for Accreditation A is 100%. So, the accuracy of the classification is 87.5%.

Keywords: Discriminant Analysis, Accreditation Score, Indicators, National education standards (SNP) early childhood, National accreditation organization (BAN)

INTRODUCTION

Accreditation is an activity carried out to determine the feasibility of educational programs and units in formal and non-formal education pathways at each level and type of education based on open criteria (RI Law Number 20 of 2003 Article 60 Paragraph (1) & (3)). These criteria can be in the form of standards as stated in Article 35 paragraph (1) which states that national standards of education consist of: content standards, process standard, graduate competency standard, education personnel standard, facilities and infrastructure standard, management standard, financing standard, and educational assessment standard which must be improved on a planned and periodic basis. In fiscal year 2016, National accreditation organization of early childhood and non-formal education (BAN PAUD dan PNF) conducted feasibility assessments of 12,500 programs from PAUD and PNF units throughout Indonesia[1].

Solution of the problem for accreditation assessment require comprehensive explanation of how to find out the maximum PAUD accreditation rating, as well as what factors affect the accreditation value of accreditation. The accreditation value factors used in this study include 8 indicators of National education standards (SNP) early childhood. 8 indicators of National education standards (SNP) early childhood included in this study are Child Development Achievement Level Standards, Content Standards, Process Standards, Educators and Education Personnel Standards, Facilities and Infrastructure Standards, Management Standards, Financing Standards, and Education Assessment Standards[2].

To classify the factors that affect the quality between one PAUD and others, and the results of accreditation in PAUD other than based on the Child Development Achievement Level Standards, The Standards of Educators and Education Personnel can use discriminant analysis [3]. Discriminant Analysis is a statistical method of grouping or classifying a number of objects into groups based on several dominant variables, such that each object becomes a member of one of the groups[4], [5]. A special feature of discriminant analysis is that the data on the bound variables is category or logistics data, while the data on the free variables is in the form of non-category data [6]. In general, discriminant analysis can be used to test accuracy and obtain classification models individually with the *Case wise Diagnostic* approach [7].

Articles related to discriminant analysis in the field of education [8] using robust quadratic discriminant analysis with Minimum Covariance Determinant (MCD) estimators to classify specialization data of SMAN 1 Kendal students containing outliers gave classification accuracy results of 95.06% with a percentage of misclassification of 4.94% while classical quadratic discriminant analysis resulted in accuracy classification of 92.59% with a percentage of misclassification of 7.41%. [9] classify the quality of high school students in each sub-district of South Aceh Regency based on final grade data. Dependent variables used are sub-district classifications and independent variables data on the average final high school / MA score for each subject tested in each department of each sub-district. The models obtained are two discriminant models for science and social studies majors. [10] classified the quality of education of 48 high schools in Lamongan Regency using discriminant analysis. The grouping of the quality of education is based on factors such as the number of classrooms, the value of accreditation, the number of certified and non-certified teachers, the number of educational personnel, the ratio of students to teachers, the number of laboratory rooms.

The results of the grouping based on the analysis of discriminant analysis obtained 58.3% of the 48 high schools in Lamongan District according to the results of school accreditation scores. [11] which applies discriminant analysis to the classification of the accuracy of the study period of Physics education students at PMIPA UPI. Variables that are seen as affecting the accuracy of the study period include basic physics scores I, GPA in the second semester, the duration of thesis preparation, PPL, the frequency of academic guidance with academic supervisors per semester, and the frequency of guidance with thesis supervisors. The results of the discriminant analysis show that the factors that distinguish timely graduation include, Basic Physics scores, the frequency of academic guidance per semester and the length of thesis preparation. The results of the analysis showed that the resulting discriminant model had an accuracy of 95.83%. [12] tested the factors influencing student learning outcomes using discriminant analysis at MI Nurul Iman, South Tangerang City. Independent variables include motivation, how to learn, teacher competence, parents environment, school infrastructure, Community Environment. The

number of respondents was 56 students. Based on the results of the discriminant analysis, it shows that the main factor that affects student learning outcomes is school infrastructure.

To our knowledge, there is no accreditation classification and mapping using discriminant analysis. Based on this fact, these articles are then used to classify PAUD in Southeast Sulawesi Province based on 8 indicators of the National education standards (SNP) early childhood and determine the accreditation score of PAUD in Southeast Sulawesi Province using discriminant analysis.

METHODS

The method used in this study used discriminant analysis. The data used in this study is secondary data on the number of accredited PAUD units in Southeast Sulawesi in 2019, as many as 697 units in 17 regencies/cities (BAN PAUD and PNF Southeast Sulawesi Province, 2019).

The bound variable used was the accreditation score for PAUD in 2019 in Southeast Sulawesi Province, which was on an ordinal scale with categories A, B, and C. Free variables in this study were 8 indicators of the National Education Standards (SNP) for Early Childhood, namely Standards for Child Development Achievement Levels (X_1), Content Standards (X_2), Process Standards (X_3), Standards for Educators and Education Personnel (X_4), Standards for Facilities and Infrastructure (X_5), Management Standards (X_6), Financing Standards (X_7), Educational Assessment Standards (X_8) [13].

The steps in the Discriminant Analysis are as follows:
1. Perform data exploration.
2. Conduct an *equality test*. To satisfy the assumption that all free variables must be equal is seen in *the table Test of Equality Group Means significance* from *Wilk's Lambda*. If the p-value > 0.05 indicates that the variable is equal. To see that the variables *are equal*, it is also seen from *the group covariance matrices* with *Box's M.* If the p-value > 0.05 means that the *covariance group* is relatively equal.
3. The significance of the discriminant function is based on the significance value of Wilk's Lambda, if the p value < 0.05, then it indicates that this discriminant function can show a clear difference between the two groups of bound variables.
4. Testing the accuracy of the classification of discriminant functions (individually).
5. Making conclusions.

RESULTS AND DISCUSSION

a. Data Exploration

Data exploration is carried out to find out the general initial information from the data. The following is shown the number of PAUD accreditation scores in Southeast Sulawesi Province in 2019 as shown in the following figure:
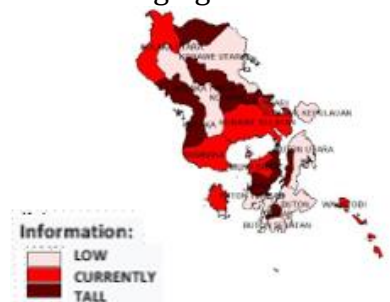

Figure 1. Mapping the Number of ECCE and PNF Units in Southeast Sulawesi Province 2019

Figure 1 shows the mapping of the number of PAUD and PNF units in Southeast Sulawesi Province in 2019. In 2019, the PAUD unit in Southeast Sulawesi Province had the largest percentage of achievements of 8 SNPs in Konawe, Muna, Kolaka, Bau-bau, and South Konawe Regencies. The percentage of achievements of 8 SNPs is in Buton, North Konawe, Buton Tengan, North Buton, Bombana, West Muna, Wakatobi, North Kolaka, and Kendari districts. In addition, the lowest percentage of 8 SNP achievements was in Konawe Islands, East Kolaka, and South Buton Districts. Figure 2 shows the profile of the results of PAUD Accreditation in Southeast Sulawesi Province in 2019.
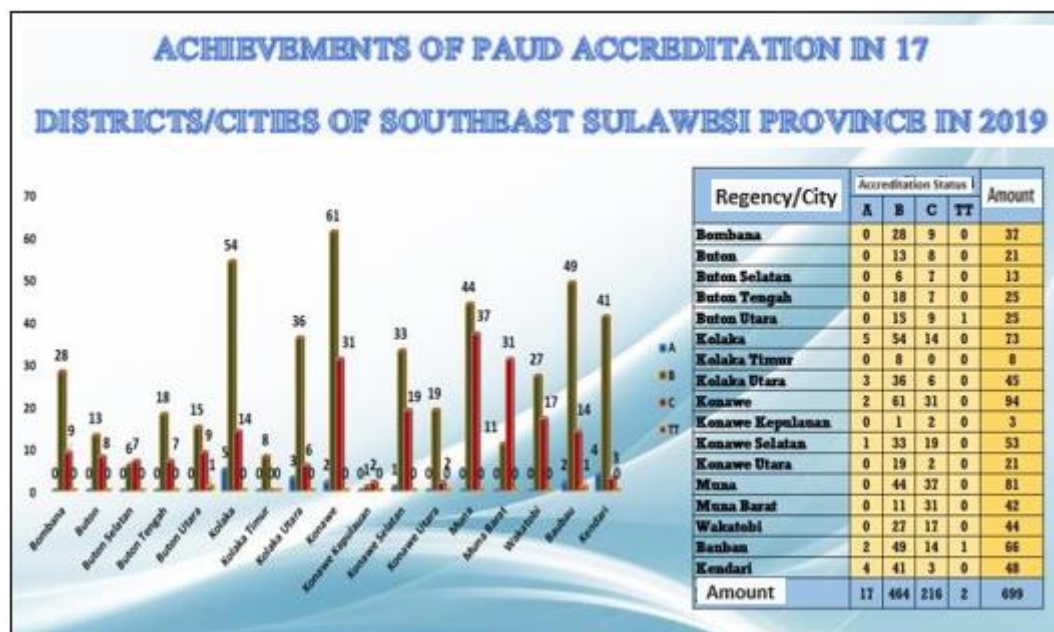

Figure 2. Profile of PAUD Accreditation Results in Southeast Sulawesi Province in 2019

b. Discriminant Analysis

The discussion in this article focuses on the issue of classifying 8 indicators of the National education standards (SNP) early childhood and determining the accreditation rating of PAUD in Southeast Sulawesi Province. Amount of data for Discriminant Analysis can see in Table 1.

Table 1. Analysis Case Processing Summary

	Unweighted Cases	N
	Valid	697
	Missing or out-of-range group codes	0
Excluded	At least one missing discriminating variabel	0
	Both missing or out-of-range group codes and at least one missing discriminating variabel	0
	Total	0
	Total	697

From the *output* of Table 1, it can be seen that the number of data processed was 697 respondents (PAUD) by entering all the data in obtained *by Group Statistics.* Based on the Case Processing analysis presented in Table 2, 697 data were declared valid. There is no data that is out of reach, indicating that the 697 PAUD data included are PAUD located in Southeast Sulawesi Province. Data for all PAUD taken also depend on the value factors

$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$, and there is no data that does not depend on at least one or more of these variables.

Table 2. *Group Statistics*

Acreditation Result		Mean	Std. Deviation	Valid N (Listwise)	
				Unweighted	Weighted
C	X_1	55030.736	14521,133	216	216
	X_2	511.574	75.362	216	216
	X_3	632.407	180.608	216	216
	X_4	57563.634	21575.119	216	216
	X_5	50724.537	35427.039	216	216
	X_6	435.185	181.628	216	216
	X_7	837.963	239.464	216	216
	X_8	780.093	253.390	216	216
B	X_1	66138.272	15198.823	464	464
	X_2	582.974	186.218	464	464
	X_3	800.000	134.694	464	464
	X_4	71848.537	18017.343	464	464
	X_5	41311.343	39259.142	464	464
	X_6	528.017	168.437	464	464
	X_7	949.353	154.556	464	464
	X_8	945.043	159.972	464	464
A	X_1	75745.294	9122.281	17	17
	X_2	911.765	196.476	17	17
	X_3	917.647	101.460	17	17
	X_4	75607.706	29278.522	17	17
	X_5	15529.352	32352.968	17	17
	X_6	764.706	257.248	17	17
	X_7	970.588	121.269	17	17
	X_8	970.588	121.268	17	17
Total	X_1	62930.368	15841.708	697	697
	X_2	568.867	172.434	697	697
	X_3	750.933	170,369	697	697
	X_4	67513.339	20598.580	697	697
	X_5	43599.659	38413.042	697	697
	X_6	505.022	184.657	697	697
	X_7	915.352	191.432	697	697
	X_8	894.549	207.611	697	697

Based on Table 2, of the 697 PAUD data used in this study, there were 216 PAUD with A accreditation, 464 with B accreditation, and 17 with C accreditation. From Table 2 it can also be seen that the average of the Standards for Child Development Achievement Levels (X1) Early Childhood Education (PAUD) in Southeast Sulawesi Province as a whole is 62930,368, with PAUD accredited A having the highest average score and PAUD accredited C having the lowest average score. The highest average score was also obtained by PAUD accredited A from Content Standards (X2), Process Standards (X3), Standards for Educators and Education Personnel (X4), Management Standards (X6), Financing Standards (X7), and Education Assessment Standards (X8)). While the highest average score of Facilities and Infrastructure Standards (X5) was obtained by PAUD with accreditation C, with a value of 50724,537 and having a difference of 35195,185 from PAUD with accreditation A which had the lowest average score. Based on the standard deviation value which is always smaller than the mean value, it indicates that there is no data deviation in the variables X_1, X_2, X_3, X_4,

$X5_1$, X_6, X_7, and X_8. Output Table 2 points to the fact that all variables X_1, X_2, X_3, X_4, $X5_1$, X_6, X_7, and X_8 are valid and reliable, so that all variables can be used for discriminant analysis.

c. Test of Discriminant Assumption
The results of *the equality test* can be seen in Table 3.

Table 3. Box Test of Equality of Covariance Matrice

Accreditation Results	Rank	Log Determinant
C	8	111.009
B	8	110.404
A	8	106.332
Pooled within-groups	8	111.193

The ranks and logarithms of natural determinants are printed group covariance matrices.
a. Rank < 2
b. Too few cases to be not singular

Table 3 shows the results of the variance similarity test, where from the table logarithm determinants it is seen that the rank for each group is 8 with the logarithm determinant for group C is 111.009, group B with logarithm determinant 110.404 and logarithm determinant group A is 106.332. The Box-M test is a multivariate statistical test used to evaluate a sample group consisting of two or more variables with normal multivariate distribution or not. The null hypothesis of Box-M states that the population mean covariance matrix with each variable is maintained at a constant number Based on Table 3, shows that groups A, B, and C have a population matrix covariance that is less than or equal to the log determinant pooled within-groups and therefore further tests can be continued about the average equation of the three-factor groups using discriminant analysis.

Table 4. Standardized Canonical Discriminant Function Coefficients

Indicator	Function	
	1	2
X_1	0.353	0.007
X_2	0.552	-0.698
X_3	0.681	0.183
X_4	0.200	0.225
X_5	-0.213	0.134
X_6	0.556	-0.338
X_7	0.606	0.222
X_8	0.633	0.325

Table 4 shows the magnitude of the contribution of the magnitude of the coefficients of the discriminator variable. *This output* shows how important the discriminator variable is in forming the discriminant function. The higher the standard value of *the Canonical Discriminant Function*, the more important it is in forming a relative variable compared to other variables. From the *output* it can be seen that the value of the *standard* coefficient for the Child Development achievement level standard is 0.353, The Content Standard is 0.552, The Process Standard is 0.681, Educator Standards 0.200, Standard of Facilities and Infrastructure -0.213, The Management Standard is 0.556, Financing Standard is 0. 606, and for the Educational Assessment Standard is 0.633. By looking at the results, it can be

concluded that the value of the eight standards can be used as a form of discriminant equation with the following results:

$$D1 = 0.353X_1 + 0.552X_2 + 0.681X_3 + 0.2X_4 - 0.213X_5 + 0.556X_6 + 0.606X_7 + 0.633X_8$$
$$D1* = 0.007X_1 - 0.698X_2 + 0.183X_3 + 0.225X_4 + 0.134X_5 - 0.338X_6 + 0.222X_7 + 0325X_8$$

Table 5. Classification Statistics

Prior Probabilities For Groups			
Accreditation Results	Prior	Cases Used in Analysis	
		Unweighted	Weighted
C	0.333	216	216
B	0.333	464	464
A	0.333	17	17
Total	1.000	697	697
Classification Processing Summary			
	Processed		697
Excluded	Missing or out-of-range group codes		0
	At least one missing discriminating variabel		0
	Used in Output		697
Classification Function Coefficients			
Indicators	Accreditation Results		
	C	B	A
X_1	0.000	0.000	0.000
X_2	0.046	0.054	0.072
X_3	0.054	0.067	0.078
X_4	0.000	0.000	0.000
X_5	-3,282E-6	-1,736E-5	-4,018E-5
X_6	0.039	0.047	0.060
X_7	0.052	0.061	0.069
X_8	0.045	0.054	0.061
(Constant)	-90.714	-129.116	-179.556

Table 5 Classification Statistics shows the total value of cases included in the 697 PAUD model, where none of the data were excluded from the analysis. The results of prior probabilities for groups showed that the probability of each group was 33% with the data entered in the analysis being 697 PAUD including 216 PAUD at value C, 464 PAUD at value B, and 17 at value A. Based on this classification, 17 PAUD with A accreditation score, 464 PAUD with B accreditation score and 216 PAUD with C accreditation score.

Then 3 classification functions were formed to divide each PAUD into A, B or C accreditation. Based on the classification results from the function coefficients, indicators X_1 and X_4 were excluded because they had no effect at all. While X_5 can be ignored because it has a very small effect. Output Classification Function Coefficients is a coefficient to form a Discriminant equation:

$$D (C) = -90.714 + 0.046 X_2 + 0.054 X_3 + 0.039 X_6 + 0.052 X_7 + 0.045 X_8$$
$$D (B) = -129,116 + 0.054X_2 + 0.067 X_3 + 0.047 X_6 + 0.061 X_7 + 0.054 X_8$$
$$D (A) = -179,556 + 0.072 X_2 + 0.078 X_3 + 0.060 X_6 + 0.069 X_7 + 0.061 X_8$$

The results of *the Classification Function Coefficients* classify the 8 indicators of the National education standards (SNP) early childhood in Southeast Sulawesi Province and predict PAUD accreditation score in Southeast Sulawesi Province. The accreditation results obtained based on the Discriminant analysis are seen in Table 6, namely: 17 units got an A, 454 units got a B and 216 units got a C (Table 6).

Table 6. *Classification Results*

Indicators		Accreditation Score	Predicted Group Membership			Total
			C	B	A	
Original	Count	C	198	18	0	216
		B	29	395	40	464
		A	0	0	17	17
	%	C	91.7	8.3	0.0	100.0
		B	6.3	85.1	8.6	100.0
		A	0.0	0.0	100.0	100.0
87.5% correct data classified						

The percentage of correctness in classifying discriminant functions resulting from Accreditation C is 91.7%, Accreditation B is 85.1%, and for Accreditation A is 100%. The above results give a classification accuracy value of 87.5%. The level of accuracy is quite high when compared to the results of studies [14]–[16] which have accuracy values ranging from 71% -78%. In that study, the resulting discriminant function depended on 2 to 3 coefficients and function variables. However, the accuracy value in this study can be increased as in research [17] which has an accuracy rate of 94.7%.

CONCLUSIONS

Based on their results of data analysis, the classification of PAUD accreditation score in Southeast Sulawesi Province was obtained as follows:

$$D (C) = -90.714 + 0.046 X_2 + 0.054 X_3 + 0.039 X_6 + 0.052 X_7 + 0.045 X_8$$
$$D (B) = -129,116 + 0.054 X_2 + 0.067 X_3 + 0.047 X_6 + 0.061 X_7 + 0.054 X_8$$
$$D (A) = -179,556 + 0.072 X_2 + 0.078 X_3 + 0.060 X_6 + 0.069 X_7 + 0.061 X_8$$

The results of uji *Wilk's Lambda* show that the variables of Content Standards (X_2), Process Standards (X_3), Management Standards (X_6), Financing Standards (X_7), and Education Assessment Standards (X_8) it is the main component in the formation of the same as the discriminant. The accreditation results obtained based on the discriminant analysis are divided into 3 classifications, namely Accreditation C is 91.7%, Accreditation B is 85.1%, and for Accreditation A is 100%. So, the accuracy of the classification is 87.5%.

REFERENCES

[1]   Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal, *Panduan Permohonan Akreditasi Program dan Satuan PAUD dan PNF Tahun 2016*. Jakarta: Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal, 2016.

[2]   I. Yuliantina, "Pentingnya Akreditasi Satuan Paud," *Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal*, 2020.

[3]   J. F. Hair Jr., W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 7th ed. New Jersey: Prentice Hall, 2010.

[4]   A. C. Rencher, *Methods of Multivariate Analysis*, 2nd Editio. New York: John Wiley & Sons, 2002.

[5] R. A. Johnson and D. W. Wichern, *Applied Multivariate Stastistical Analysis*. New Jersey: Prentice Hall, 2007.

[6] Wahana, *Pengembangan Analisis Multivariat dengan SPSS 12*. Jakarta: Salemba, 2005.

[7] J. Supranto, *Analisis Multivariat Arti dan Intepretasi*. Jakarta: Rineka Cipta, 2004.

[8] L. i. N. Khiqmah, M. A. Mukid, and A. Prahutama, "Perbandingan Diskriminan Kuadratik Klasik dan Diskriminan Kuadratik Robust Pada Kasus Pengklasifikasian Peminatan Peserta Didik (Studi Kasus di SMA Negeri 1 Kendal Tahun Ajaran 2014/2015)," *J. Gaussian*, vol. 4, no. 2, pp. 295–304, 2015.

[9] A. R. Fitriana, Nurhasanah, and R. R. Jannah, "Klasifikasi Kecamatan Berdasarkan Nilai Akhir SMA/MA di Kabupaten Aceh Selatan Menggunakan Analisis Diskriminan," *Statistika*, vol. 15, no. 2, pp. 87–97, 2015.

[10] M. Billah, N. E. Chandra, and S. Amiroch, "Analisis Diskriminan Terhadap Pengelompokan Mutu Pendidikan Sekolah Menengah Atas di Kabupaten Lamongan," *Unisda J. Math. Comput. Sci.*, vol. 6, no. 2, pp. 11–20, 2020.

[11] D. Saepuzaman and H. Retnawati, "Analisis Disriminan Dan Klasifikasi Ketepatan Masa Studi Mahasiswa Pendidikan Fisika," *J. Pendidik. Fis. Univ. Islam Negeri Alauddin Makasar*, vol. 9, no. 2, pp. 92–102, 2021.

[12] Sulaeman, "Faktor Yang Mempengaruhi Hasil Belajar Siswa Menggunakan Analisis Diskriminan MI Nurul Iman Kota Tangerang Selatan," *Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 2, no. 3, pp. 256–266, 2021.

[13] Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal Provinsi Sullawesi Selatan, *Pedoman Kebijakan dan Mekanisme Akreditasi PAUD dan PNF Provinsi*. Kendari: Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal Provinsi Sullawesi Selatan, 2018.

[14] T. Ramayah, N. H. Ahmad, H. A. Halim, S. Rohaida, M. Zainal, and M. Lo, "Discriminant analysis : An illustrated example," *African J. Bus. Manag.*, vol. 4, no. 9, pp. 1654–1667, 2010.

[15] D. D. K. Ghosh and S. S. Kulshrestha, "Comparison of Employee Branding Through Discriminant Analysis," *Int. J. Manag.*, vol. 7, no. 7, pp. 395–405, 2016, [Online]. Available: http://iaeme.com/Home/issue/IJM?Volume=7&Issue=7

[16] G. M. D. Ganga, K. F. Esposto, and D. Braatz, "Application of discriminant analysis-based model for prediction of risk of low back disorders due to workplace design in industrial jobs," *Work*, vol. 41, no. SUPPL.1, pp. 2370–2376, 2012, doi: 10.3233/WOR-2012-0467-2370.

[17] D. Sokari and O. Emmanuel, "Discriminant Analysis of Unemployment and Literacy Rate by State in Nigeria," vol. 3, no. 1, pp. 44–60, 2020.