IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.16, No.3, July 2022, pp. 301~312
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: https://doi.org/10.22146/ijccs.70433

Implementation of Factor Analysis and BiClustering in Classifying Multidimensional Under-Five Poverty in East Nusa Tenggara

Rahmadathul Wisdawati*[1], Rani Nooraeni[2], Bagaskoro Cahyo Laksono[3], Bintang Izzatul Fatah[4]
[1]BPS-Statistics of Karimun Regency, Indonesia
[2,3,4]Polytechnic of Statistics STIS, East Jakarta, Indonesia
e-mail: ***[1]rahma.wisdawati@bps.go.id**, [2]raninoor@stis.ac.id, [3]211810211@stis.ac.id, [4]211810215@stis.ac.id

Abstrak

Kemiskinan balita merupakan keadaan tidak tercukupinya semua kebutuhan dasar balita sehingga mengakibatkan balita kekurangan gizi dan tidak dapat mencapai potensi diri secara penuh dalam lingkup sosial. Provinsi Nusa tenggara Timur merupakan provinsi yang masih menghadapi tantangan dalam permasalahan gizi tertinggi di Indonesia pada tahun 2019. Penelitian ini bertujuan untuk menggambarkan kondisi kemiskinan multidimensi balita di NTT, membentuk Indeks Kemiskinan Balita Multidimensi (IKBM), serta membandingkan hasil pengelompokan kabupaten/kota di NTT berdasarkan indeks yang terbentuk dengan pengelompokan kabupaten/kota berdasarkan metode biclustering. Sumber data yang digunakan pada penelitian ini adalah SUSENAS KOR 2019, Metode analisis yang digunakan adalah analisis faktor dan bicluster. Hasil penelitian menunjukkan bahwa dari 11 indikator kemiskinan multidimensi terbentuk 3 dimensi yaitu Faktor Fasilitas Makanan dan Minuman Layak (FFMML), Faktor Perlindungan Kesehatan (FPK) dan Faktor Perumahan dan Nutrisi (FPN) yang digunakan untuk membentuk indeks. Menurut pengelompokan wilayah, terdapat lima wilayah berkategori nilai IKBM rendah, empat belas wilayah berkategori nilai IKBM sedang, dan tiga wilayah berkategori tinggi untuk nilai IKBM. Sedangkan berdasarkan hasil biclustering, terdapat dua wilayah berkategori kemiskinan rendah, tiga belas wilayah berkategori kemiskinan sedang dan tujuh wilayah berkategori kemiskinan tinggi. Hasil perbandingan pengelompokan IKBM dengan metode biclustering diperoleh hasil yang berbeda dilihat dari komposisi wilayah yang dihasilkan.

***Kata kunci**—Kemiskinan Balita, Analisis Faktor, BiClustering*

Abstract

Under-five poverty is a condition where the needs of toodlers are not met, resulting in undernourished children and unable to reach their full potential in the social sphere. East Nusa Tenggara is a province that still faces the biggest nutritional problems in Indonesia in 2019. This study aims to explain the variables that form toodlers multidimensional poverty in East Nusa Tenggara (ENT), form the Multidimensional Under-Five Poverty Index (MUPI), and compare the results of index formed with the results of bicluster. Data source used in this study is SUSENAS KOR 2019. The analytical method used is a factor and bicluster analysis. The results shows that 11 multidimensional poverty indicators form three dimensions, namely the Adequate Food and Beverage Facility Factor, Health Protection Factor, and Housing and Nutrition Factor, which is used to form the index. Based on regional grouping, there are five areas with low MUPI scores, fourteen areas with medium MUPI scores, and three areas with high MUPI scores. However, biclustering results show that there are two areas with low poverty category, thirteen regions with moderate poverty category, and seven regions with high

Received November 14th,2021; Revised July 14th, 2022; Accepted July 31th, 2022

poverty category. The result of the comparison of MUPI grouping with the biclustering method obtained different results based on the composition of the resulting area.

Keywords—*Under-five poverty, factor Analysis, BiClustering*

1. INTRODUCTION

Poverty and inequality are real problems for every country, especially for developing countries. Child poverty is one of the many challenges that society must face together. After the financial crisis in Asia, the poor population in Indonesia in 2010 was 13.3 percent which in 2018 decreased to 9.82 percent [1]. Despite experiencing a downward trend, millions of people, including children, still fall into poverty.

When the concept of child poverty expanded, the poverty does not only accounted by income earned by households. The poverty also includes the living conditions of a child. For example, if access to health, education, water and sanitation, food, and nutrition were accounted, it would appear that nine out of ten Indonesian children are still affected by poverty in one or more critical dimensions of child welfare. The child poverty is a multidimensional problem because many factors cause children to become poor [2].

Apart from the children's living conditions, other factors clarify the disparity between groups of children, such as significant inter-provincial disparities. The poverty rate in DKI Jakarta is the lowest at 3.8 percent. However, poverty in Papua has a big gap with DKI Jakarta, which is 27.8 percent (BPS, 2018). Those Provinces in eastern Indonesia are more significantly affected by poverty compared to the western and central regions, where the average difference between the five poorest provinces in the eastern region is 18 percent higher than the five wealthiest provinces in Indonesia [3].

East Nusa Tenggara is the province with the lowest nutritional status in Indonesia. Based on the publication of the 2019 Indonesian health profile published by the Ministry of Health of the Republic of Indonesia (Kemenkes RI), East Nusa Tenggara has the third-highest percentage of poor people in Indonesia (21.09 percent), which is quite far from the percentage of poor people nationally (9,41 percent). Based on the nutritional status of children under five, East Nusa Tenggara has the highest percentage of malnutrition and malnutrition in children under five, which are 6.9 percent and 17.6 percent, respectively. Malnutrition is calculated based on the bodyweight index based on age (BB/U) which is stipulated in the Decree of the Minister of Health Number 1995/Menkes/SK/XII/2010. The malnourished children will have an impact on decreased immunity so that they are vulnerable to infectious diseases [4].

In 2017, the Central Bureau of Statistics conducted a study entitled "Analysis of Child Poverty and Deprivation of Children's Basic Rights in Indonesia", analyzing the poverty level and characteristics of children in Indonesia with a monetary approach and a multidimensional approach with the Multiple Overlapping Deprivation Analysis (MODA) method. The method build six dimensions, namely housing, facilities, food and nutrition, education, child protection, and health. The concept of children are divided into two, namely the age of 0-4 years (toddlers) and 5-17 years. The result of the analysis show that 64.95 percent of children aged 0-17 years are multidimensionally deprived, where the dimension that contributes the largest in the 0-4 year age group is the health dimension of 76.50 percent, while in the 5-17 year age group the most deprived in the age group.

In its development, other studies discuss multidimensional poverty, such as in the research of Lalu and Ika [5], which forms a multidimensional poverty index in children aged 0-17 years in Indonesia using Multiple Overlapping Deprivation Analysis (MODA). Rahmadatul and Rani [6] built a poverty index with a multidimensional approach for children under five in NTT Province, using factor analysis as a method for building the poverty index. On the other hand, Christiana, Irfani, and Sartono [7] used bicluster analysis of the Cheng and Church (CC) algorithm based on poverty dimensions to find patterns of poverty characteristics in Central Java Province. The biclustering method can group each observation based on the same

characterizing variable so that the grouping results can map the area based on the similarity of the characteristics of each variable [8].

Based on the problems and studies that has described previously, there has been no further research that compares the result of multidimensional grouping poverty in children under five with the formation of an index and biclustering, so some of the objectives of this study are as follows: to describe the variables forming the multidimensional poverty of children under five in East Nusa Tenggara, to form the Multidimensional Under-Five Poverty Index (MUPI) based on variables forming multidimensional under five poverty, grouping areas based on MUPI and using the biclustering method, and comparing the results of grouping areas based on MUPI and biclustering. The results of this analysis will produce suggestions for better grouping methods in explaining multidimensional under-five poverty so that the policies provided can be more targeted in alleviating the poverty under five in Indonesia.

2. METHODS

2.1 Under-Five (Toddler) Poverty

In general, poverty is a condition in which a person or group of people, men, and women, are not fulfilled their fundamental rights. For example like food, clothing, and health properly to develop and lead a dignified life that departs from a rights-based approach that recognizes the same fundamental rights as other members of society. Toddler poverty is a deprivation situation (a psychological state of feeling dissatisfaction or gaps or deficiencies) experienced by toddlers. Toddlers are individuals or groups of individuals from a population who are within a spesific age range. The age of toddlers under five grouped into three groups, namely the infant age group (0-2 years), the toddler group (2-3 years), and the preschool group (>3-5 years). Meanwhile, based on WHO, the age group of toddlers is 1-5 years [9].

2.2 Multidimensional Under-Five Poverty Index (MUPI)

Multidimensional poverty of toddlers and children is known as MODA (Multiple Overlapping Deprivation Analysis). MODA is a comprehensive measure of the multidimensional level of deprivation or non-fulfillment of children's fundamental rights. In compiling the Multidimensional Toddler Poverty Index, the various dimensions used are the housing dimension, the facility dimension, the food and nutrition dimension, the education dimension, the housing dimension, and the health dimension. The data used in this study is data from the 2019 KOR National Socio-Economic Survey (SUSENAS) in East Nusa Tenggara to form MUPI. Based on the previously described dimensions, the indicators used in this study are as follows:

Table 1 Indicators Used in Research

No	Indicators	Description
1	Floor Area (X_1)	The percentage of toddlers who live in a house with a floor area per capita < 7.2 m^2.
2	Floor Type (X_2)	The percentage of toddlers who live in the house with the broadest type of floor comes from the ground.
3	DrinkingWater (X_3)	The percentage of toddlers who do not have access to adequate drinking water sources.
4	Sanitation (X_4)	The percentage of toddlers who do not have access to proper sanitation.
5	Cooking Fuel (X_5)	The percentage of toddlers who are in the household using natural fuels (charcoal, wood, and the like) as the primary fuel for cooking.
6	Breastfeeding (X_6)	The percentage of toddlers who do not receive exclusive breastfeeding (if they are less than six months old) and do not receive continued breastfeeding for up to 24 months (for ages 0-23 months)
7	Early Initiation of Breast Milk (X_7)	The percentage of toddlers who are not placed on the mother's chest or do not do IMD, and when placed on the mother's chest but the process of attaching the baby is less than 1 (one) hour.
8	Education (X_8)	The percentage of toddlers aged 3-4 years not attending preschool education
9	Birth Certificate (X_9)	The percentage of toddlers who do not have a birth certificate.
10	Immunization(X_{10})	The percentage of toddlers who did not receive complete immunizations
11	Health Insurance (X_{11})	The percentage of toddlers who do not have health insurance.

2.3 Factor Analysis

Factor analysis is a multivariate statistical method to explain the diversity of data structures. The results of the factor analysis will group variables into factor, which is used for the subsequent analysis. Based on Johnson and Wichern [10], the model in factor analysis is as follows:

$$(X-\mu)_{(p\times 1)} = L_{(p\times m)}F_{(m\times 1)} + \varepsilon_{(p\times 1)} \quad (1)$$

One of the estimation methods used in factor analysis is the Principal Component Method. The form of the equation is as follows:

$$\Sigma = \lambda_1 e_1 e_1^{'} + \cdots + \lambda_p e_p e_p^{'} = \left[\sqrt{\lambda_1}e_1 \vdots \cdots \vdots \sqrt{\lambda_p}e_p\right]\left[\sqrt{\lambda_1}e_1^{'} \ \vdots \ \sqrt{\lambda_p}e_p^{'}\right] = LL^{'} + 0 = LL' \quad (2)$$

It is appropriate to use the covariance matrix if the observed variable has a size on a scale that is not much different or has the same units. However, if the variables have different units, standardization is necessary by:

$$z_j = \begin{bmatrix} \frac{x_1-\bar{x}_1}{\sqrt{s_{11}}} \\ \frac{x_2-\bar{x}_2}{\sqrt{s_{22}}} \\ \vdots \\ \frac{x_j-\bar{x}_j}{\sqrt{s_{pp}}} \end{bmatrix} ; j = 1, 2, \ldots, n \quad (3)$$

Information:

X_i = observed variable value to-*i* / *j*; *i* = 1, 2, … , *p*.
μ_i = mean of i-variable
ε_i = i-th *specific factor*
F_j = *j*th *common factor*; j = 1, 2, … ,*m*
l_{ij} = *loading* of the ith variable on the *j*th factor
p = many variables observed
m = the number of common factors formed
L = factor *loading* matrix
Z_j = the standardization value of the variable to-*j*
$\bar{X}_j$ = the average of the variables to-*j*
s = standard deviation

2.4 Bi Clustering

In this research, the biclustering analysis used is the CC (Cheng & Church) or - biclustering algorithm. The biclustering analysis is clustering carried out in two directions to find row subgroups and column subgroups that have a high correlation [11]. Clustering is a data mining technique used to form clusters containing objects that have similarities and between clusters that have differences [12]. The CC algorithm is a greedy algorithm, which tries to find the maximum bicluster with high similarity. A set of rows and columns are called as bicluster if it has a mean squared residue value below the level (δ) that determined by the researcher [13]. The following is the calculation of the mean squared residue:

$$H(I,J) = \frac{1}{|I||J|}\Sigma_{i\epsilon I, j\epsilon J}\left(a_{ij} - a_{iJ} - a_{Ij} + a_{IJ}\right)^2 \quad (4)$$

Then, to compare the quality of the bicluster results from several different, the average bicluster ratio, which is the ratio between residue and volume, is used as follows [14].

$$\frac{1}{b}\sum_{i=1}^{b}\frac{Residue_i}{Volume_i} \quad (5)$$

$$Volume_i = c_i \times d_i \frac{1}{b}\sum_{i=1}^{b}\frac{Residue_i}{Volume_i} \quad (6)$$

Through this ratio, the selected biclustering is the one with the smallest ratio value. In addition, the quality of the resulting bicluster is also assessed for coherence in the following value of additive model variance [15].

$$b_{ij} = \pi + \beta_i + \beta_j \quad (7)$$

which π is the value of the matrix element, while β_i $(1 \le i \le |I|)$ and β_j $(1 \le j \le |J|)$ show constant values in the additive model for each row and column. If the value of the additive variance is zero, then the bicluster is ideally coherent. Generally, a value above 1.5 is

sufficient to determine that the bicluster is incoherent. The CC biclustering method has used in several studies. In Yuniarto & Kurniawan's research [16], this method used to map the dimensional structure of poverty in East Java Province. Then Azizah [17] used this method to determine groups of active compounds from medicinal plants that are similar to synthetic drugs and have a relationship with protein groups that play a role in type 2 diabetes. Industry 4.0. In addition, Kaban et al. [19] also use this method to map districts/cities in Indonesia based on the variables that make up the social vulnerability index by Siagian et al. [20].

2.5 MANOVA

Based on Johnson and Wichern [10], the model in manova is as follows:

$$X_{lj} = \mu + \tau_l + e_{lj,} j = 1,2,\dots,n_l, and\ l = 1,2,\dots,g \quad (8)$$

Where the e_{lj} are independent $N_p(0,\Sigma)$ variables. Here the parameter vector μ is an overall mean (level), and τ_l represents the lth treatment effect with $\sum_{l=1}^{g} \quad n_l\tau_l = 0$. The hypothesis is:

$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$

$H_1: at\ least\ two\ \mu's\ are\ unequal$

3. RESULTS AND DISCUSSION

3.1 Overview of The Poverty Condition of Under-Five in East Nusa Tenggara

The condition of the multidimensional poverty indicator in every district/city in East Nusa Tenggara can be visible in Figure 1. The further away from the central point, the more basic rights are not fulfilled in the region based on the indicators. Early initiation of breast milk is an indicator with the highest level of deprivation, while the lowest level of deprivation is the floor area. Some regions that have low indicator scores are Kupang city, Malaka, and Southwest Sumba. Whereas the regions that have high indicator scores are Lembata, Nagekeo, and Sikka.

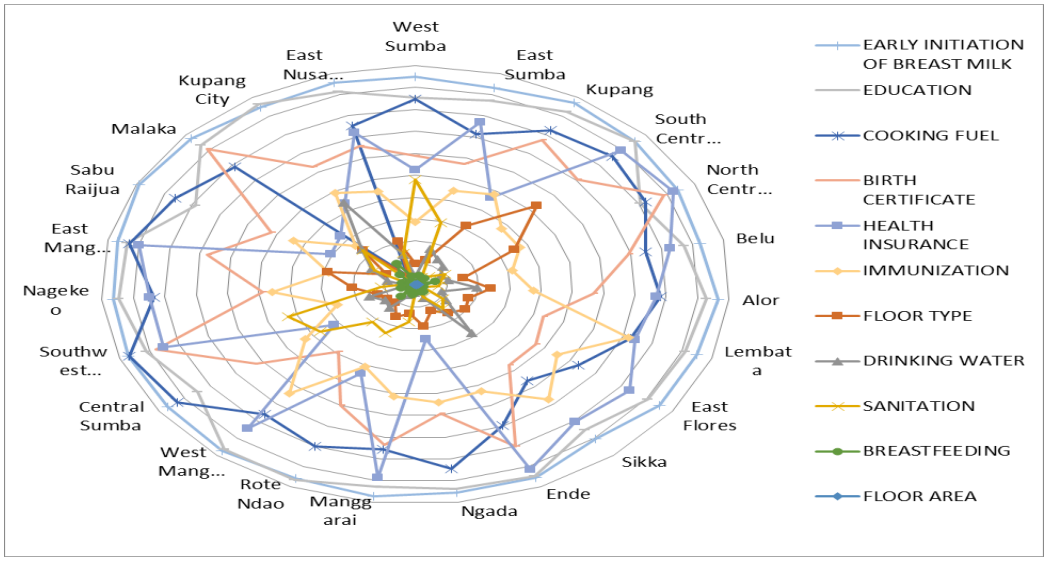

Figure 1 Poverty Condition based on MODA with 11 Indicators

3.2 Forming and Grouping of the Multidimensional Under-Five Poverty Index (MUPI)

The development of the Multidimensional Under-Five Poverty Index in each district/city in East Nusa Tenggara Province in 2019 in this study uses 11 selected indicators. First, standardize all these indicators using equation (3). Then, using factor analysis in equation (1) and equation (2) was implemented on the data.

Table 2. The result of the KMO, Bartlett, and MSA test for each phase for validating an Indicator of MUPI in East Nusa Tenggara

Test	Indicators	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
KMO		0,435	0,477	0,532	0,603	0,638
Bartlett		0,001	0,000	0,000	0,001	0,000
MSA	X_1	0,307	0,315	(Reducted)	(Reducted)	(Reducted)
	X_2	0,340	0,415	0,374	0,569	0,617
	X_3	0,423	0,534	0,576	0,643	0,665
	X_4	0,449	0,389	0,339	(Reducted)	(Reducted)
	X_5	0,559	0,612	0,587	0,659	0,645
	X_6	0,443	0,475	0,677	0,662	0,755
	X_7	0,425	0,400	0,522	0,608	0,654
	X_8	0,431	0,467	0,612	0,569	0,561
	X_9	0,259	(Reducted)	(Reducted)	(Reducted)	(Reducted)
	X_{10}	0,599	0,587	0,627	0,547	0,542
	X_{11}	0,522	0,448	0,406	0,361	(Reducted)

Phase 1 includes 11 indicators for the developing of a multidimensional under-five poverty index. The data is correlated with the indicators used if the p-value of the Bartlett test <0.05. The data is feasible for factor analysis if the KMO value is at least 0.5 [21]. The declared variable can be used for further analysis if the MSA value is at least 0.5.

Finally, in Phase 5, there are seven indicators that will be analyzed; further that is, type of floor, drinking water, cooking fuel, breastfeeding, early initiation of breast milk, birth certificates, and immunizations. The Measures of Sampling Adequacy (MSA) in Table 2 shows the *p-value* in the Bartlett test is 0.000 and the KMO at stage five is 0.638. It means that the indicator is worthy of further analysis in factor analysis.

Table 3 Loading values of Rotated Component Matrix and Grouping Variables in East Nusa Tenggara

No	Indicators	component			Factor
		1	2	3	
1	Type of floor	-0,413	0,478	0,565	Factor 3
2	Drinking water	0,878	0,0024	-0,110	Factor 1
3	Cooking Fuel	0,840	-0,331	0,210	Factor 1
4	Breastfeeding	0,880	0,135	0,085	Factor 1
5	Early initiation of Breast Milk	-0,261	0,229	0,807	Factor 3
6	Birth Certificate	0,064	0,890	-0,029	Factor 2
7	Immunization	-0,060	0,856	-0,181	Factor 2

Furthermore, look at the results of the characteristic root value (eigenvalue) and the percentage of variance explained to determine the number of factors. The grouping of each indicator into three factors can be seen from the closeness of the loading value. in order to get results that are easy to interpret, then do the rotation using the varimax method.

Table 3 shows that the loading values in the rotated component matrix, each indicator is highly correlated with only one factor. The first factor has three indicators that are proper drinking water, cooking fuel, and breastfeeding. The second factor has two indicators, that are birth certificates and complete immunization. The third factor has two indicators, which are the type of floor and early initiation of breast milk. The name of the first factor is Adequate Food and Beverage Facility (AFBF). The second factor's name is Health Protection Factor (HPF). The name of the third factor is Housing and Nutrition Factor (HNF). The weight used in this research is unequal weight. The use of this weight is because there are differences in the variance described in each factor [21]. Table 4 shows the weight value of each factor.

Table 4 Weighted Factor Scores of MUPI in East Nusa Tenggara

Factor	%Described Variance	Weight
AFBF	35,706	0,4544
HPF	27,630	0,3516
HNF	15,251	0,1941
Total	78,587	1

The weight of AFBF, HPF, and HNF are 0,4544; 0,3516; 0,1941 consecutively. Thus, the equation of MUPI is as follows :

$$MUPI = (0{,}4544 \times AFBF)_i + (0{,}3516 \times HPF)_i + (0{,}1941 \times HNF)_i \quad (8)$$

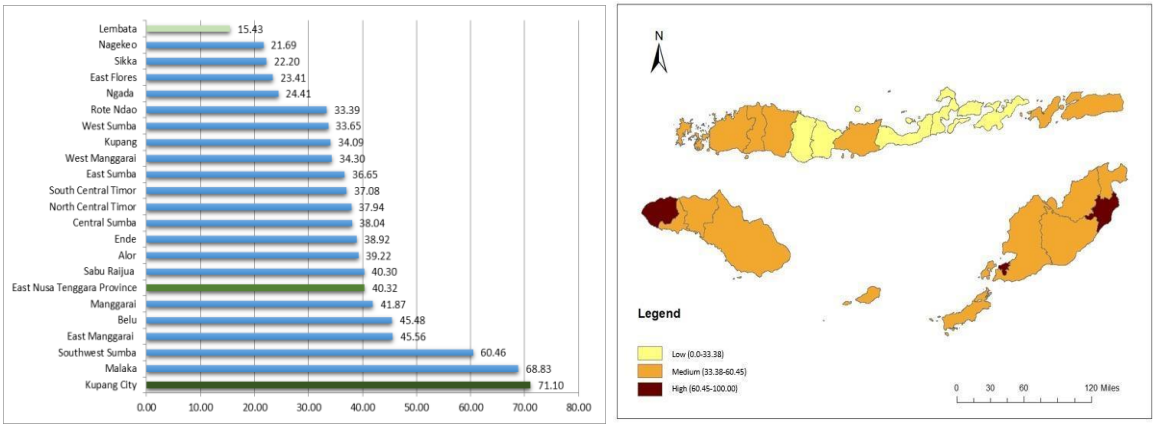

Figure 2 Histogram of MUPI and its group based on natural break in East Nusa Tenggara

Figure 2 shows the distribution of MUPI scores in East Nusa Tenggara. Lower scores show a better condition for Under-five poverty. The Value of MUPI scores in East Nusa Tenggara is 40,32 percent. Sixteen districts have scored smaller than the province scores, and six districts have scored higher than the province scores. The district with the best MUPI score is the Lembata district, which is 15,43 percent. That means the level of deprivation in Lembata is low. The worst MUPI score is Kota Kupang with 71,10 percent. This means that many toddlers still lack their rights in various dimensions of poverty and welfare in Kota Kupang.

There are five cities with low MUPI score, that is Lembata, Ngada, Nagekeo, Sikka, and Flores Timur. 14 districts/cities with medium MUPI score are Alor, Ende, Manggarai, Manggarai Barat, Sumba Barat, Kupang, Timor Tengah Selatan, Timor Tengah Utara, Sabu Raijua, Sumba Timur, Belu, Sumba Tengah, Rote Ndao, and Manggarai Timur. The other cities with high MUPI score are Malaka, Sumba Barat Daya, and Kota Kupang.

Table 5 One-way MANOVA Result based on total number of Indicators

Number Indicators	Pillai	Df	App F	Num df	Den df	p-value
all indicators (11)	2	1,244	1,4958	22	20	0,1846
selected indicators (7)	2	1,0589	2,2505	14	27	0,03292

Results of the test, One-way MANOVA (Table 5), show that there are different characteristics of each form group based on the selected indicators (7 indicators). The indicators that can differentiate the area are: the percentage of children under five who live in the house with the broadest type of floor comes from the ground, the percentage of children under five who do not have access to adequate drinking water sources, the percentage of children under five who are in the household using natural fuels (charcoal, wood, and the like) as the main fuel for cooking, the percentage of toddlers who do not receive exclusive breastfeeding (if they are less than six months old) and do not receive continued breastfeeding for up to 24 months (for ages 0-23 months), the percentage of toddlers who are not placed on the mother's chest or do not do IMD, and also when placed on the mother's chest but the process of attaching the baby is less

than 1 (one) hour, percentage of toddlers aged 3-4 years not attending preschool education and the percentage of children under five who do not have a birth certificate.

Table 6 Simultaneous Confidence Interval MUPI Group

Simultaneous Confidence Interval						
Indicators	Low and medium MUPI		Low and high MUPI		Medium and high MUPI	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
Floor Area	-0.309197	0.2238558	-0.305849	0.4413603	-0.215044	0.4358979
Floor Type	-13.56074	10.37375	-11.46164	22.08861	-7.706925	21.52088
Drinking Water	-8.923702	8.586188	-29.41765	-4.873104	-27.6678	-6.285444
Sanitation	-23.72274	5.681562	-34.96531	6.252273	-23.28958	12.61773
Cooking Fuel	-28.3521	10.4475	-15.2303	39.15716	-2.774501	44.60596
Breastfeeding	-3.43952	1.119898	-8.611028	-2.21985	-7.039514	-1.471743
Early Initiation of Breast Milk	-3.462063	0.6975174	-4.885024	0.9456815	-3.127152	1.952355
Birth Certificate	-29.32831	-3.139853	-49.80165	-13.0919	-31.20282	0.7774334
Education	-4.780851	6.782631	-10.65157	5.55758	-10.60831	3.512537
Immunization	5.367206	28.17475	8.904404	40.87496	-5.807113	22.04452
Health Insurance	-26.86349	26.23409	-21.47916	52.95055	-16.36988	48.47068

3.3 Grouping district/cities in East Nusa Tenggara based on Toddler poverty indicators with Bi Clustering

We group all regions in East Nusa Tenggara with the bicluster method in all indicators. There are three bicluster and nine characterized variables obtained that is immunization, education, birth certificate, early initiation of breast milk, breastfeeding, cooking fuel, drinking water, floor type, and floor area.

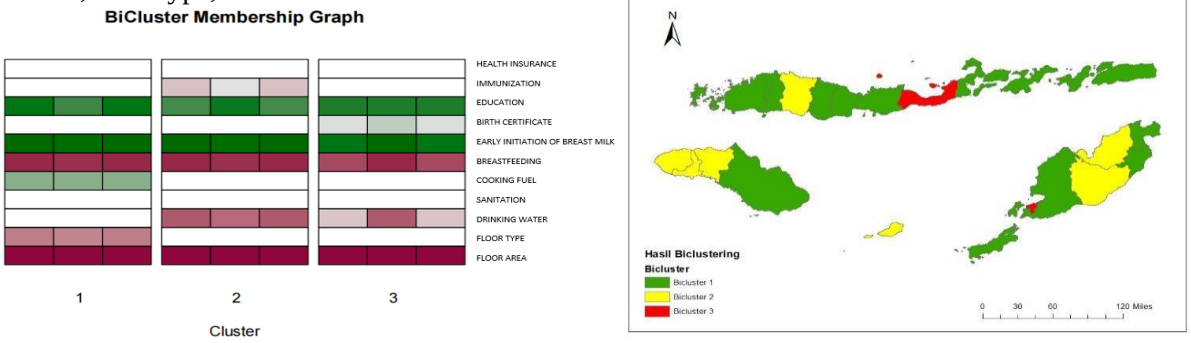

Figure 3. The BiCluster Membership Graph based on Characterized Variables and Thematic Map Bicluster Group

Bicluster 1 could group 13 district/cities, that are Sumba Timur, Kupang, Belu, Alor, Lembara, Flores Timur, Ende, Ngada, Rote Ndao, Manggarai Barat, Nagekeo, and Malaka. The variables that characterized the district/cities are floor area, type of floor, cooking fuel, breastfeeding, early initiation of breast milk, and education. Bicluster 2 can group 7 districts/cities that are Sumba Barat, Timor Tengah Utara, Timur Tengah Selatan, Sumba Tengah, Sumba Barat Daya, Manggarai Timur, Sabu Raijua. The six characterizing variables are floor area, drinking water, breastfeeding, early initiation of breast milk, birth certificate, and immunization. The last bicluster contains Sikka and Kupang City with floor area, drinking

water, breastfeeding, early initiation of breast milk, birth certificate, and education as characterizing variables. We also conduct the MANOVA from biclusters.

Table 7 One-way MANOVA Result based on total number of Indicators

Number Indicators	Pillai	Df	App F	Num df	Den df	p-value
all indicators (11)	2	1,3609	2,839	18	24	0,009007
selected indicators (9)	2	1,4079	2,1618	22	20	0,04401

Results of the test, One-way MANOVA (Table 7), show that there are different characteristics of each form group based on the selected indicators (9 indicators). The indicators that able to differentiate the area are : Percentage of children under five who live in a house with a floor area per capita < 7.2 m^2, the percentage of children under five who live in the house with the broadest type of floor comes from the ground, the percentage of children under five who do not have access to adequate drinking water sources, the percentage of children under five who are in the household using natural fuels (charcoal, wood, and the like) as the primary fuel for cooking, the percentage of toddlers who do not receive exclusive breastfeeding (if they are less than six months old) and do not receive continued breastfeeding for up to 24 months (for ages 0-23 months), the percentage of toddlers who are not placed on the mother's chest or do not do IMD, and also when placed on the mother's chest but the process of attaching the baby is less than 1 (one) hour, the percentage of toddlers aged 3-4 years not attending preschool education, the percentage of children under five who do not have a birth certificate, and the percentage of children under five who did not receive complete immunizations.

Table 8 Simultaneous Confidence Interval of Bi Cluster Group

Simultaneous Confidence Interval						
Indicators	Bicluster 1 and 2		Bicluster 1 and 3		Bicluster 2 and 3	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
Floor Area	-0.261924	0.2139705	-0.309589	0.4546558	-0.304634	0.497655
Floor Type	-15.43802	5.692166	-9.34284	24.59035	-5.31453	30.3079
Drinking Water	-6.785413	6.06007	-35.82962	-15.20093	-35.98041	-14.3248
Sanitation	-22.21408	3.476782	-14.3224	26.93481	-5.980674	37.33038
Cooking Fuel	-22.33227	-2.353536	29.84122	61.92527	41.38553	75.06677
Breastfeeding	-1.887887	2.781516	-8.015759	-0.517117	-8.649219	-0.777285
Early Initiation of Breast Milk	-2.103912	1.315309	0.5460348	6.036997	0.8036626	6.567972
Birth Certificate	-20.6482	7.729442	-14.57119	31.00075	-9.246132	38.59444
Education	0.8845929	10.32164	-5.504405	9.650646	-11.48474	4.424748
Immunization	1.562128	20.55383	-27.39518	3.103793	-39.2123	-7.195048
Health Insurance	-24.02602	25.61626	-33.70072	46.02032	-36.48015	47.20951

3.4 Comparing the group of districts based on toddler poverty indicators between groping of MUPI score and Biclustering

The characteristics of each group area that resulted from MUPI and Bicluster can be described based on the summary values of each indicator (Table 9). Based on the mean statistics, the result concludes that the order of poverty from lowest to highest is Bicluster 3 (Low), Bicluster 1 (Medium), and Bicluster 2 (High), based on cooking fuel. For MUPI, the order of categories is appropriate, namely Group 1 (Low), Group 2 (Medium), and Group 3 (High) based on drinking water, breastfeeding, and birth certificate.

Table 9 Statistical Summary of every Indicator MUPI in ENT Province

Indicators	Summary	Bicluster			MUPI		
		Cluster 1	Cluster 2	Cluster 3	Grup 1	Grup 2	Grup 3
Floor Area	Min	0	0	0	0	0	0
	Median	0	0	0	0	0	0
	Mean	0.07253	0.09651	0	0.6776	0.1104	0
	max	0.8704	0.67557	0	0.33878	0.8704	0
Type of Floor	Min	10.52	9.62	4.325	16.51	9.62	4.325
	median	17.47	12.58	10.42	19.05	14.24	12.579
	mean	18.04	22.92	10.42	18.61	20.2	13.292
	max	31.37	52.83	16.514	20.35	52.83	22.971
Drinking Water	Min	0.88	3.974	28.49	0.8854	3.797	15.79
	median	11.3268	11.858	36.23	8.9047	11.2	23.54
	mean	10.7187	11.081	36.23	10.6258	10.795	27.77
	max	23.5434	15.905	43.98	28.4882	19.788	43.98
Sanitation	Min	1.1	2.098	0.7015	1.109	2.098	0.7016
	median	10.884	14.644	7.3203	10.748	12.663	22.8342
	mean	13.627	22.995	7.3203	7.821	16.841	22.1771
	max	29.299	47.886	13.939	13.939	47.886	42.9956
Cooking Fuel	Min	63.26	82.53	4.913	56.34	69.7	4.913
	median	77.11	86.22	30.628	72.64	81.38	78.555
	mean	76.51	88.85	30.628	72.03	80.98	60.062
	max	84.46	96.72	56.342	84.46	93.1	96.719
Breastfeeding	Min	0	0.8612	4.109	0	0	4.889
	median	2.984	2.0672	7.729	2.547	3.07	6.601
	mean	3.462	2.734	7.729	2.197	3.357	7.613
	max	6.933	4.8886	11.348	4.109	6.933	11.348
Early Initiation of Breast Milk	Min	92.54	93.83	90.46	90.46	92.54	94.86
	median	96.24	96.52	92.61	95.19	96.24	97.12
	mean	95.9	96.3	92.61	94.54	95.92	96.51
	max	97.66	99.18	94.76	96.58	99.18	97.66
Birth Certificate	Min	39.16	51.77	46.97	42.94	39.16	62.99
	median	59.52	67.71	54.98	46.97	64.96	87.6
	mean	63.2	69.66	54.98	48.97	65.2	80.41
	max	90.65	89.07	62.99	59.28	89.07	90.65
Education	Min	87.12	78.79	85.13	85.13	78.79	91.01
	median	93.06	85.3	90.65	90.95	92.55	93.34
	mean	92.72	87.12	90.65	90.96	89.96	93.51
	max	96.93	95.68	96.18	94.61	96.93	96.18
Immunization	Min	25.95	26.55	49.12	45.82	28.05	25.95
	median	47.07	37.15	58.08	5.05	41.69	26.55
	mean	45.93	34.87	58.08	58.77	41.99	33.88
	max	63.85	43.64	67.03	71.91	63.85	49.12
Health Insurance	Min	24.82	30.21	43.56	24.82	30.21	32.66
	median	79.84	85.15	61.96	80.36	79.84	43.56
	mean	68.12	67.33	61.96	69.53	69.84	53.79
	max	91.83	92.33	80.36	84.97	92.33	85.15

Based on the cross-tabulation between the categorization with the MUPI and the bicluster, ten regions are categorized the same as seen from the diagonal of the matrix. For areas with different categorization results due to differences in the characteristics of the indicators used to build the bicluster and MUPI. Based on the results, it can be concluded that the results of the bicluster with the results of the MUPI cannot be said to be in harmony. Therefore, it can be considered that biclustering is insufficient to describe the poverty condition of children under five in East Nusa Tenggara. The difference in classification occurs because the Floor Area indicator is a marker for the formation of each cluster but not an indicator for forming MUPI.

Table 10 Cross-tabulation Bicluster group with Natural Break MUPI group

		MUPI		
		Grup 1 (Low)	Grup 2 (Medium)	Grup 3 (High)
Bi Cluster	Cluster 3 (Low)	Sikka	0	Kupang City
	Cluster 1 (Medium)	Lembata, Nagekeo, East Flores, Ngada	East Sumba, Kupang, Belu, Alor, Ende, Rote Ndao, West Manggarai	Malaka
	Cluster 2 (High)	0	West Sumba, North Timor Tengah, South Timor Tengah, Sumba Tengah, Barat Daya, East Manggarai, Sabu Raijua.	Southwest Sumba

4. CONCLUSIONS

Factor analysis using the principal component extraction method resulted in three factors identifying under-five poverty in East Nusa Tenggara. Based on the results of Manova, the grouping based on the Natural Breaks MUPI scheme proved to be different in each group formed based on seven sets of indicators forming the index, namely Floor Type, Drinking Water, Cooking Fuel, Breastfeeding, Early Initiation of Breast Milk, Birth Certificate and Immunization. Based on the results of biclustering, nine indicators were obtained, namely Floor Type, Floor Area, Education, Drinking Water, Cooking Fuel, Breastfeeding, Early Initiation of Breast Milk, Birth Certificate, and Immunization. The use of the biclustering method as a comparison against grouping based on the Natural Breaks MUPI scheme produces different results seen from the tabulation of each result. In this case, the grouping of the biclustering method is not as complex as the grouping of the results of the MUPI. However, the results of the biclustering cannot be said to be better because the results are different from the formation of the MUPI. In this research, researchers still feel there are shortcomings when selecting the number of clusters that have not used the optimum parameters. Suggestions for future research are to use better cluster selection criteria.

REFERENCES

[1] UNICEF. (2020). *Situasi Anak di Indonesia - Tren, Peluang, dan Tantangan dalam Memenuhi Hak-hak Anak.* Jakarta: UNICEF Indonesia.

[2] Badan Pusat Statistik. (2017). *Analisis Kemiskinan Anak dan Deprivasi Hak-hak Dasar Anak di Indonesia*. Jakarta: Badan Pusat Statistik.

[3] OECD. (2019). *Social Protection System Review of Indonesia, OECD Development Pathways, OECD.* Paris: OECD Publishing.

[4] Alamsyah, D., Mexitalia, M., Margawati, A., Hadisaputro, S., & Setyawan, H. (2017). Beberapa Faktor Risiko Gizi Kurang dan Gizi Buruk pada Balita 12-59 Bulan (Studi Kasus di Kota Pontianak). *Jurnal Epidemiologi Kesehatan Komunitas*, *2*(1), 46. https://doi.org/10.14710/jekk.v2i1.3994

[5] Aulia, L. A., & Wulansari, I. Y. (2020). Pembentukan Indeks Kemiskinan Multidimensi Anak Dan Pemanfaatannya Untuk Pengentasan Kemiskinan Berkelanjutan Di Indonesia Tahun 2017. *Seminar Nasional Official Statistics*, *2019*(1), 336–346. https://doi.org/10.34123/semnasoffstat.v2019i1.222

[6] Wisdawati, R., & Nooraeni, R. (2021). *Pembangunan Indeks Kemiskinan Balita Multidimensi di Provinsi Nusa Tenggara Timur Tahun 2019*. *March 2019*.

[7] Putri, C. A., Irfani, R., & Sartono, B. (2021). Recognizing poverty pattern in Central Java using Biclustering Analysis. *Journal of Physics: Conference Series*, *1863*(1). https://doi.org/10.1088/1742-6596/1863/1/012068

[8] Nurmawiya., & Kurniawan, R. (2021). Pengelompokan Wilayah Indonesia Dalam Menghadapi Revolusi Industri 4.0 Dengan Metode Biclustering. *Seminar Nasional Official Statistics*, *2020*(1), 790–797. https://doi.org/10.34123/semnasoffstat.v2020i1.511

[9] Andriani, Merrryana dan Wirjatmadi, Bambang. (2012). *Peranan Gizi dalam Silkus Kehidupan*. Jakarta: Kencana.

[10] Johnson, Richard A and Wichern, Dean W. (2007). *Applied Multivariate Statistical Analysis*. New Jersey : Pearson Education, Inc.

[11] Sundari, M., Sihombing, P. R., & Hakim, K. F. (2021). Perbandingan Metode Analisis Gerombol K-Rataan Dan Bicluster (Studi Kasus: Kerentanan Kelurahan Di Kota Depok TAHUN 2020). *Lombok Journal of Science*, 1-11.

[12] Habibi, M., & Cahyo, P. W. (2019). Clustering User Characteristics Based on the influence of Hashtags on the Instagram Platform. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *13*(4), 399. https://doi.org/10.22146/ijccs.50574

[13] Mina, E., Feelders, A. J., Kemmeren, P., & Siebes, A. P. J. M. (2011). Applying biclustering to understand the molecular basis of phenotypic diversity. *Phd. Utrecht University Faculty of Science Department of Information and Computing Sciences*.

[14] Chakraborty, A., & Maka, H. (2005 Biclustering of gene expression data using genetic algorithm. In *2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology* (pp. 1-8). IEEE.

[15] Pontes, B., Girldez, R., & Aguilar-Ruiz, J. S. (2015). Quality measures for gene expression biclusters. PloSone, 10(3)

[16] Yuniarto, B., & Kurniawan, R. (2017). Understanding Structure of Poverty Dimensions in East Java: Bicluster Approach. *Signifikan: Jurnal Ilmu Ekonomi*, *6*(2), 289-300.

[17] Rifai, N. A. K., Afendi, F. M., & Sumertajaya, I. M. (2018). Simultaneous clustering analysis with molecular docking in network pharmacology for type 2 antidiabetic compounds. *Indonesian Journal of Biotechnology*, *22*(1), 43. https://doi.org/10.22146/ijbiotech.27334

[18] Nurmawiya., & Kurniawan, R. (2021). Pengelompokan Wilayah Indonesia Dalam Menghadapi Revolusi Industri 4.0 Dengan Metode Biclustering. *Seminar Nasional Official Statistics*, *2020*(1), 790–797. https://doi.org/10.34123/semnasoffstat.v2020i1.511

[19] Kaban, P. A., Kurniawan, R., Caraka, R. E., Pardamean, B., & Yuniarto, B. (2019). Biclustering Method to Capture the Spatial Pattern and to Identify the Causes of Social Vulnerability in Indonesia: A New Recommendation for Disaster Mitigation Policy. *Procedia Computer Science*, *157*, 31-37.

[20] Siagian, T. H., Purhadi, P., Suhartono, S., & Ritonga, H. (2014). Social vulnerability to natural hazards in Indonesia: driving factors and policy implications. Natural hazards, 70(2), 1603-1617.

[21] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*, 7th *Edition*. United State: Pearson Prentice Hall.