

The Implementation of Biplot and Cluster Analysis to Construct Sustainable Education Policies Toward Technology Literate Society in Indonesia

Pressylia Aluisina Putri Widyangga¹

¹Faculty of Science and Technology, Universitas Airlangga, Indonesia
Jl. Dr. Ir. H. Soekarno, Mulyorejo 60115, Indonesia, Telp 0315915551

Corresponding author

pressylia.aluisina.putri-2021@fst.unair.ac.id

Abstract: The digital era encourages society, especially the younger generation, to have the ability in developing technology and adapt in the era full of change and uncertainty. To establish a technology literate society in Indonesia, it is necessary to enhance the quality of education. This research aims to analyse the correlation between the competence in using technology, literacy rates, and the education completion level in establishing a technology literate society. Through biplot analysis, it was found that the education completion level, literacy rate, and the percentage of people who are able to use technology has a high positive correlation. Therefore, provinces where society has the capability to develop technology are provinces that have high literacy rates and educational completion level. In addition, using cluster analysis with the average linkage method, provinces in Indonesia are grouped into four clusters based on the erudition regarding technology, where the first cluster consists of five provinces, second cluster with 22 provinces, third cluster with six provinces, and fourth cluster with a province, namely Papua. Through biplot and cluster analysis, the government can construct sustainable policies that are right on target based on the problems faced by each province.

Keywords: Biplot, cluster, digital, Indonesia, technology.

Introduction

The digital era is a condition where all activities that support life can be facilitated by the presence of sophisticated technology. This term can also be interpreted as the emergence of digital technology that replaces the technologies which were previously used. Society is a person or group of people who form a semi-closed or semi-open system where most of the interactions are between individuals who are in the group. The development of society is divided into five phases, including society 1.0, society 2.0, society 3.0, society 4.0, and society 5.0 (Sandi, 2020). During society 1.0, humans began to learn about forming a group to become a society and defend themselves by hunting for food and moving places to survive. Society 2.0 is referred to the era of agriculture or the agricultural revolution where humans focus on developing science because of farming. Society 3.0

is an era where people are increasingly focused on farming accompanied by an increasingly complex number of people, making the need for food and clothing increase. Society 4.0 is an era where science and technology are increasingly developing, making people familiar with computers and the internet so they can acquire information rapidly. Society 5.0 is a refinement of society 4.0, where technology becomes a part of the humans, not only for sharing information, but for facilitating human life.

The digital era requires people in every country, including Indonesia, to have the competence in developing technology. To realize this, it is necessary to increase the quality of education so that people, especially the younger generation, have the skill to develop technology through the education received. Knowledge is a factor that has the aim of capturing whether a country has the ability, awareness, understanding of something

needed to learn and discover new technologies. Technology is a factor that has the intention of assessing the readiness of a country to enter the development phase of digital transformation. The main problem is that not all people in Indonesia have the ability to adapt in the digital era. The majority of Indonesia citizen do not have knowledge related to current technological developments and innovations. Based on Badan Pusat Statistik (BPS), only 70.2% of Indonesian people have competence in using Information and Communication Technology in 2021. This figure is an increase when compared to the percentages in 2019 and 2020 (BPS, 2021). However, the percentage of people who have skills in using Information and Communication Technology must be increased. This is a challenge for Indonesia, seeing that there are still limited access and improper quality of education in Indonesia.

To establish an Indonesian society that has the competence in developing technology, research is needed to find out what factors are behind the level of society erudition of technological developments. Therefore, a study will be carried out related to the application of biplot and hierarchical cluster analysis to overcome existing problems. The biplot analysis used in this study will analyse the correlation between related variables, such as literacy rates, technology literacy rates, and educational completion rates at the elementary, junior high and senior high school levels. Biplot analysis can assist in analysing whether these factors can affect the level of public knowledge of technology. Hierarchical cluster analysis using the average linkage method will classify provinces in Indonesia based on the erudition of the society in the province about technological developments. This research is expected to assist the government in constructing policies related to education in order to create a technology-literate Indonesian society and young generation

Literature Review

Biplot Analysis

Biplot is a two-dimensional representation of the \mathbf{Y} matrix data showing each point from n observation vectors (rows from \mathbf{Y}) as well as variables (columns

of \mathbf{Y}). According to Sartono (2003), there are four pieces of information regarding the appearance of the biplot, including the following.

1. **Characteristic**
Proximity between observed objects is objects that have similar characteristics with other objects. If two points are depicted with adjacent positions, then the two objects are said to have the same characteristics.
2. **Variable diversity**
Variables that have large variability values are drawn with long vectors while variables that have small variability values are depicted with short vectors.
3. **Correlation**
The correlation between variables shows how a variable affects or is influenced by other variables. Variables will be described as directed lines. Two variables that have a positive correlation value (+) are depicted as two lines with the same direction or forming a narrow angle ($< 90^\circ$). Two variables that have a negative correlation value are depicted as two lines in opposite directions or forming a wide angle ($> 90^\circ$). Two uncorrelated variables are depicted as lines with angles approaching 90° .
4. **An object is said to have a value above the average if the object is located in the direction of the variable vector. If the object is located opposite the direction of the variable vector, then the object is said to have a value below the average. If the object is almost in the middle, then the object is said to have a value that is close to the average.**

Main component $z = \mathbf{a}'\mathbf{y}$, where are the eigenvectors of \mathbf{S} , the sample covariance matrix, and \mathbf{y} are the observed vectors. There are p eigenvectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$ so that there are z_1, z_2, \dots, z_p principal components for each observation \mathbf{y}_i for $i = 1, 2, \dots, n$. By using the central model, the observation vectors can be transformed into $z_{ij} = \mathbf{a}'_j(\mathbf{y}_i - \bar{\mathbf{y}}) = (\mathbf{y}_i - \bar{\mathbf{y}})' \mathbf{a}_j$ for $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$. Each $p \times 1$ observation vector is

transformed into $p \times 1$ principal component vector where $\mathbf{z}'_i = (\mathbf{y}_i - \bar{\mathbf{y}})'(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p) = (\mathbf{y}_i - \bar{\mathbf{y}})'\mathbf{A}$. Matrix \mathbf{A} is a matrix of size $p \times p$ whose columns are eigenvectors of \mathbf{S} that have been normalized. Principal components can also be written in the following equation.

$$\mathbf{Z} = \mathbf{Y}_c \mathbf{A} \tag{1}$$

$$\mathbf{z} = \begin{pmatrix} z'_1 \\ z'_2 \\ \vdots \\ z'_n \end{pmatrix}, \quad \mathbf{Y}_c = \begin{pmatrix} (\mathbf{y}_1 - \bar{\mathbf{y}})' \\ (\mathbf{y}_2 - \bar{\mathbf{y}})' \\ \vdots \\ (\mathbf{y}_n - \bar{\mathbf{y}})' \end{pmatrix} \tag{2}$$

An exact two-dimensional representation of \mathbf{Y}_c is obtained with the first two columns taken from and the first two columns from \mathbf{A} . If the resulting matrix is denoted by \mathbf{Z}_2 and \mathbf{A}_2 , the following equation will be obtained.

$$\mathbf{Y}_c \cong \mathbf{Z}_2 \mathbf{A}_2' \tag{3}$$

The adequacy of equation (3) can be evaluated by examining the first two eigenvalues of matrix \mathbf{S} namely λ_1 and λ_2 .

$$\frac{\lambda_1 + \lambda_2}{\sum_{i=1}^p \lambda_i} \tag{4}$$

Equation (4) shows that \mathbf{Y}_c provides a good visual representation of the plot. Matrix $\mathbf{Y}_c = \mathbf{Z}\mathbf{A}'$ can be constructed with a single value decomposition shown by the equation below.

$$\mathbf{Y}_c = \mathbf{U}\mathbf{\Lambda}\mathbf{V}' \tag{5}$$

Where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ is a diagonal matrix with square root elements of non-zero eigenvalues $\lambda_1^2, \lambda_2^2, \dots, \lambda_p^2$ of $\mathbf{Y}_c'\mathbf{Y}_c$ and $\mathbf{Y}_c\mathbf{Y}_c'$, the column of \mathbf{U} is the corresponding eigenvector of $\mathbf{Y}_c\mathbf{Y}_c'$, and the column of \mathbf{V} is the corresponding eigenvector of $\mathbf{Y}_c'\mathbf{Y}_c$.

The product $\mathbf{U}\mathbf{\Lambda}$ in the equation is equal to \mathbf{Z} , the principal component matrix scores in the equation (1). This equation can be proven by multiplying by \mathbf{V} , which is orthogonal because the matrix contains eigenvectors that have been normalized from the symmetric matrix.

$$\mathbf{Y}_c \mathbf{V} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}'\mathbf{V} = \mathbf{U}\mathbf{\Lambda} \tag{6}$$

Here is a two-dimensional representation of the matrix \mathbf{Y}_c based on main component and the same representation as the equation on a single value decomposition.

$$\mathbf{Y}_c \cong \mathbf{Z}_2 \mathbf{A}'_2 = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \\ \vdots & \vdots \\ z_{n1} & z_{n2} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \dots & a_{p1} \\ a_{12} & a_{22} & \dots & a_{p2} \end{pmatrix} \tag{7}$$

Thus, each observation is represented as a linear combination and the coordinates become elements of vector (z_{i1}, z_{i2}) and the axes become elements of vector (a_{j1}, a_{j2}) . Therefore, it is necessary to plot the dot (z_{i1}, z_{i2}) , $i = 1, 2, \dots, n$, and dot (a_{j1}, a_{j2}) , $j = 1, 2, \dots, p$. To see the difference and show the relationship of the points to the axes, the dots (a_{j1}, a_{j2}) are connected to the origin by straight lines forming arrows.

The Euclidean distance between two points (z_{i1}, z_{i2}) and (z_{k1}, z_{k2}) is approximately equal to the distance between the corresponding points (rows) \mathbf{y}'_i and \mathbf{y}'_k and in the \mathbf{Y} data matrix. The cosine of the angle between arrows or lines drawn to each pair of axes points (a_{j1}, a_{j2}) and (a_{k1}, a_{k2}) indicates the correlation between two variables. Thus, a small angle between two vectors that is less than indicates that the two variables are positively or high correlated. Conversely, the angle between the two vectors that is greater than 90° indicates that the variable has a negative correlation.

Cluster Analysis

Cluster analysis is a type of multivariate method that has the main goal of grouping objects based on their characteristics. Cluster analysis classifies objects so that every object has similarities with other objects in the same cluster. The clusters formed have high internal homogeneity and external heterogeneity. There are three measures to quantify the similarity between objects, namely correlation, distance, and association. The procedure applied must be able to group objects that have a high similarity into the same cluster.

The concept of similarity is fundamental in cluster analysis. There are three methods that can be applied, namely correlation, distance, and association.

i. Correlation

This measure can be applied to data with a metric scale, but is rarely used because the point is concerned with the value of a certain pattern, whereas the weight of cluster analysis is the size of the object. The similarity between objects can be seen from the correlation coefficient between pairs of objects as measured by several variables.

ii. Distance

The distance is the most frequently used and is applied to metric-scale data. The short distance means that an object is more similar to another object. The difference with the correlation measure is that clusters based on the size of the correlation may not have the same value but have the same pattern, while clusters based on the size of the distance have more similar values even though the patterns are different.

iii. Association

The association measure is used to measure nonmetric scale data, such as nominal or ordinal.

Cluster analysis has two methods, namely hierarchical and non-hierarchical methods. The hierarchical method begins by grouping two or more objects that have the closest similarity. Then, the process is continued to another object that has a second proximity. So, the cluster will form a kind of "tree", where there is a clear hierarchy between objects, from the most similar to the least similar. Logically, all objects will eventually form a cluster. Dendrogram is usually used to help clarify the hierarchical process.

In the hierarchical cluster method, there are two basic types, namely agglomerative and divisive. In the agglomerative method, there are five well-known methods, namely Single Linkage, Complete Linkage, Average Linkage, Ward's Method, and Centroid Method. Single Linkage is based on the smallest distance, which means that if two objects are separated by a short distance, then the two objects will be merged into one cluster. Complete Linkage is the opposite method to Single

Linkage because the grouping is based on the furthest distance. Average Linkage is a procedure that is almost the same as Single Linkage or Complete Linkage, but the criteria used are the average distance of all individuals in one cluster to the distance of all individuals in other clusters.

The average linkage method is a clustering method with the principle of the average distance between each possible pair of objects in one cluster with all objects in another cluster. Average linkage calculates the distance between two clusters which is referred to the average distance for each cluster. The average linkage procedure starts by defining the matrix $\mathbf{D} = \{d_{ik}\}$ to obtain the closest objects, for example U and V. These objects are combined into clusters (UV) then the distance between (UV) and other clusters, (W).

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{UV}N_W} \quad (8)$$

d_{ik} = distance between object i in cluster (UV) and object k in cluster W

N_{UV} = number of items in UV cluster

N_W = number of items in W cluster

Materials and Methods

Data Sources and Variables

The data used in this research is secondary data obtained from the Badan Pusat Statistik (BPS). The variables in this study include the proportion of people who have the competence in using technology, literacy rates, and educational completion rates at the elementary, junior high, and senior high school levels in 2021. Table 1 provides an explanation of each variable definition used in this study.

Table 1. Research Variables

Variables	Definition
Proportion of Society Which Have Competence in Using Technology	Proportion of youth and adults aged 15-59 who have skills in using information and communication technology in the digital era.
Literacy Rate	Proportion of population aged 15 years and over who has the ability to read and write simple sentences in letters. Literacy rate is an important indicator to see how far the society of an area are open to knowledge.
Education Completion Rate	Residents who have completed education at a certain level after being registered and actively participating in education, both formal education (elementary, junior high, and senior high school) and non-formal which are under the supervision of the Ministry of National Education, Ministry of Religion, other government and private agencies.

Data Analysis Procedure

The following explanation is the data analysis procedure used in this study so that the research objectives can be achieved.

1. Obtaining secondary data from Badan Pusat Statistik (BPS) regarding the proportion of people who have technology development skills, literacy rates, and education completion rates at the elementary, junior high, and senior high school levels by province in Indonesia in 2021.
2. Summarizing the data in tabular form to simplify the process of data analysis.
3. Conducting biplot analysis to determine the correlation between variables in the study.
4. Conducting a cluster analysis to classify provinces in Indonesia based on technology literacy so that the government can construct appropriate education policies based on the problems faced by each province.
5. Making conclusions and suggestions that are expected to assist the government in constructing policies for the increasing the number of technologically literate people.

Results and Discussion

Descriptive Statistics

Table 2. Descriptive Statistics

Variable	Mean	Standard Deviation	Data Size
Technology Literacy Rate	66.8344	11.40063	34
Literacy Rate	96.3985	4.18370	34
Elementary School Completion Rate	96.3671	3.69944	34
Junior High School Completion Rate	88.3226	5.84927	34
Senior High School Completion Rate	65.5106	10.51026	34

Based on the descriptive statistics table on Table 2, it can be seen that the literacy rate in Indonesia is

very high because it is more than 95% or close to 100% (BPS, 2021). This figure means that almost all Indonesian people can read and write so there is a great possibility in capturing information and knowledge.

Majority of Indonesian people have completed education at the elementary school as evidenced by the average completion rate of education at the elementary school of more than 95% (BPS, 2021). The level of completion of junior high school education in Indonesian society is very high because it is at almost 90%. However, this number still needs to be increased so that the Indonesian people can have proper knowledge regarding the technology developments. In addition, the average level of completion of Indonesian public education at the senior high school level is relatively low because it is at 65%. This means that majority of Indonesian people do not complete the 12-year compulsory education program. The government needs to pay more attention to overcome these problems so that the Indonesian people, especially the younger generation, have proper erudition regarding developments that are taking place in the current digital era, especially related to technological developments.

Low level of completion of Indonesian society's education at the senior high school level is followed by the low proportion of Indonesian people who have competence in using information and communication technology. Thus, it can be seen that the education will greatly affect people's knowledge regarding technological developments. In order for Indonesian people, especially the younger generation, to have capability in using technology and knowledge regarding technological developments, they must complete the 12-year compulsory education program and involve technology in the education process.

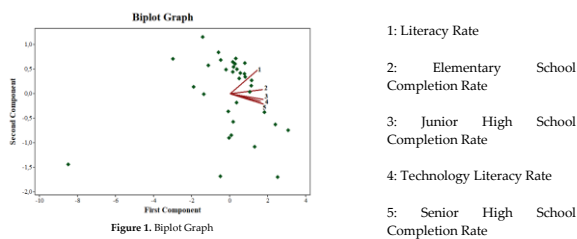
Biplot Variance

Table 3. Biplot Variance

Component	Total	Eigenvalues % of Variance	Eigenvalues Cumulative %
1	3.748	74.957	74.957
2	0.567	11.334	86.291
3	0.409	8.181	94.471
4	0.188	3.760	98.232
5	0.088	1.768	100.000

Based on Table 3, the variance that can be explained by all types of natural disasters is 100%. However, by conducting a biplot analysis and reducing all factors that influence the society erudition regarding technology development into a two-dimensional space, only 86.291% of the information that can be explained which is obtained from the sum of 79.957% and 11.334%. Therefore, it can be said that the biplot provides sufficient information regarding the relationship between the five variables.

Biplot Analysis



Based on Figure 1, it can be seen that the five research variables, such as literacy rate, technology literacy rate, and education completion level at the elementary school, junior high school, and senior high school have a high correlation because the angles between the vectors are close. In addition, the five variables have a positive correlation because they are located at an angle of less than 90 degrees. This statement implies that the level of public knowledge of technology is influenced by the community's ability to read and write as well as the education that the community has taken. The higher the level of education an individual has taken, it will be easier for the individual to acquire new knowledge, adapt in the era of disruption, and the higher level of individual erudition regarding technological developments.

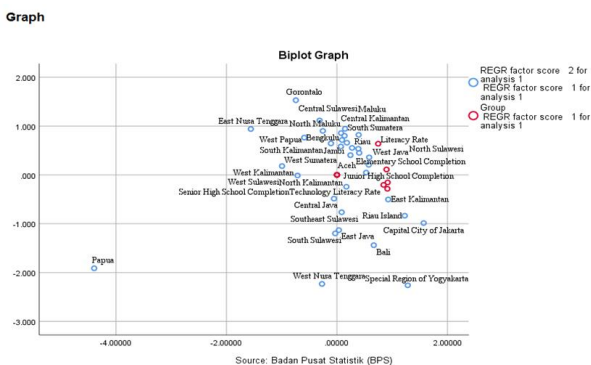


Figure 1. Biplot Graph

Figure 2 shows the correlation of the five variables and the provinces in Indonesia based on their characteristics. For example, Papua does not have a tendency towards these five variables, so the Indonesian government must pay more attention to improve the quality and access to education in Papua so that people in that province are able to compete with people in other provinces and can have the capability and knowledge in developing technology. Graph in Figure 2 is closely related to provincial groupings based on community acquaintance of technology which will be discussed using hierarchical cluster analysis using the average linkage method.

Hierarchical Cluster Analysis

The following figure is a dendrogram that groups provinces in Indonesia based on the education that the community has attained and the level of public knowledge regarding technology development, followed by a table showing the names of the provinces in each cluster. Each different colour on the dendrogram indicates that the provinces are in the different cluster. The hierarchical cluster analysis used in this study was divided provinces in Indonesia into four groups, including provinces whose the society have felicitous, well, moderate, and poor levels of education and knowledge of technology.

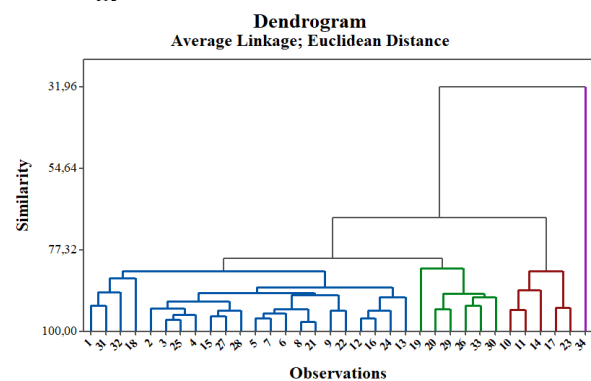


Figure 2. Hierarchical Cluster Analysis Dendrogram

Table 4. Provinces in Each Cluster

Cluster	Provinces
1	Riau Island (10), Capital City of Jakarta (11), Special Region of Yogyakarta (14), Bali (17), East Kalimantan (23)
2	Aceh (1), North Sumatera (2), West Sumatera (3), Riau (4), Jambi (5), South Sumatera (6), Bengkulu (7), Lampung (8), Bangka Belitung (9), West Java (12), Central Java (13), East Java (15), Banten (16), West Nusa Tenggara (18), Central Kalimantan (21), South Kalimantan (22), North Kalimantan (24), North Sulawesi (25), South Sulawesi (27), Southeast Sulawesi (28), Maluku (31), North Maluku (32)
3	East Nusa Tenggara (19), West Kalimantan (20), Central Sulawesi (26), Gorontalo (29), West Sulawesi (30), West Papua (33)
4	Papua (34)

Explanation of Biplot and Cluster Analysis

Based on Figure 2, it can be seen the tendency of provinces in Indonesia towards the five research variables such as literacy rate, technology literacy rate, and education completion level at the elementary school, junior high school, and senior high school. The provinces in the first cluster, namely Riau Island, Capital City of Jakarta, Special Region of Yogyakarta, Bali, and East Kalimantan, are in positions that are quite close to the five variables. This means that society in the provinces in the first cluster have high literacy rates, levels of educational completion, and the ability in developing technology.

The provinces in the second cluster are in a position that is further from these five variables or is in a position that is close to only a few variables. For example, the province of Aceh is located close to the variable education completion rate at the elementary and junior high school levels. This means that the majority of people in Aceh have completed their education at the elementary and junior high school levels. However, there are still many people in Aceh who have not completed education up to the senior high school so that Aceh is included in the province in the second cluster. Besides Aceh, Riau is also included in the second cluster. Riau's position in Figure 2 is close to the literacy rate and elementary school completion rate variables, which means that the level of people who have the capability to read and write and have completed elementary school education is relatively high. In addition, Riau's position is farther away from the junior and senior high school completion rates and technology literacy rates. This shows that there are still many people in Riau who have not completed their education up to the junior and senior high school levels and the ability of the society in Riau in developing

technology still needs to be improved. For this reason, Riau belongs to the second cluster.

Gorontalo is a province with a farther location with the five variables compared to Aceh. In Gorontalo, community education completion rates at the elementary and junior high school levels are very high. However, the completion rate of community education at the senior high school level, literacy rate, and community competence to develop technology are still low, so that Gorontalo is in the third cluster. Apart from Gorontalo, East Nusa Tenggara is also included in the third cluster and is located far away with literacy rate, technology literacy rate, elementary school, junior high school, and senior high school completion rate. This is because the majority of people in East Nusa Tenggara already have the ability to read and write and have completed their education at the elementary school level. However, people who have completed their education at the junior and senior high school levels and have the aptitude in using technology are still relatively low.

In addition, Papua is in the farthest position from the five variables so that Papua is classified in fourth cluster based on the level of education and erudition of technology. This statement means that society in Papua have low levels of capability to read and write, low education completion rate, and do not have proper skills in using technology. The government should pay special attention to Papua so that people in the province get suitable access and proper quality of education.

Sustainable Education Policies

Policies compiled by the government must be in accordance with the weaknesses of the provinces in each cluster. Sustainable policies that are on target will be able to restore the condition of the Indonesian people more rapidly till the Indonesian people have innovation, knowledge, and the ability to adapt in qualified technology. Figure 4 is created with the aim of providing solutions that the government can apply to provinces in Indonesia according to hierarchical clusters using the average linkage approach.

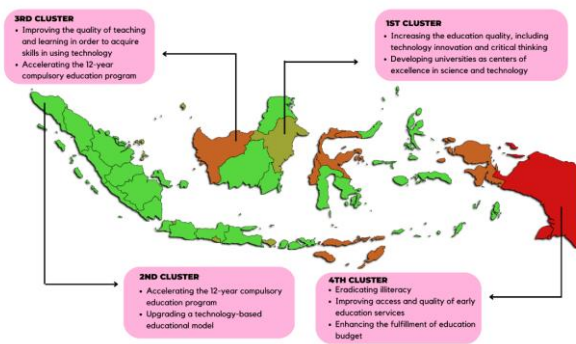


Figure 3. Suggested Solutions for Each Cluster

Figure 4 is based on the results of hierarchical cluster analysis as a visual illustration to determine solutions and policies for each province in each cluster. The suggested solution is based on the needs in each cluster in order that the solution created can establish the technology literate Indonesian people and enhance the education quality in Indonesia. Communities in the province in the first cluster already have competent reading and writing skills, high level of educational completion, and the ability to develop advanced technology. For this reason, it will be better for the government to maintain the access and quality of education in the provinces in the first cluster such as developing the quality of university as further education so that people, especially the younger generation, can acquire new erudition related to technological development.

Provinces in the second and third clusters have almost the same problems, namely the low level of educational completion and the community's ability in using technology. To solve these problems, it would be better if the government increased the access to education in order to complete the 12-year compulsory education program. If there are people who experience economic difficulties, the government can provide educational assistance funds or scholarships so that the community can complete their education properly. In addition, the use of technology needs to be involved in the learning process in both formal and non-formal education till people make technology as the part of their lives so that people have high capabilities in developing technology. For example, computer or internet can be used in the learning process, exams are carried out in a

computer-based test, and the use of online classes in collecting various assignments.

Communities in the province in the fourth cluster, Papua, have low levels of reading and writing skills. Thus, it would be better if the government first eradicated illiteracy by providing basic education, facilitating access to education, and improving the quality of education through educators and a supportive environment. The government can also provide educational assistance through scholarships to help less fortunate people to get proper education. With the existence of solutions according to the problems in each cluster, the quality of education and the ability of the community to develop technology in Indonesia can be enhanced

Conclusions

Based on the biplot analysis, it can be concluded that literacy rate, technology literacy rate, elementary school, junior high school, and senior high school completion rate have high correlation, which means that the level of education received by the community determines whether the community easily accepts new erudition and has the ability in developing technology or not. In addition, hierarchical cluster analysis using the average linkage method classifies provinces in Indonesia into four groups based on the level of education and knowledge of the community regarding technology development. Based on the analysis results in each cluster, it is known that the majority of provinces in Indonesia have the competence to read and write and have completed their education at the elementary school level. However, there are still many Indonesian people who have not completed their education at the junior and senior high school levels and do not have the capability in developing technology. This needs to be addressed because the level of education an individual receives in society determines whether the individual is open to new knowledge and can adapt in the digital era. The government must immediately construct a sustainable education policy in Indonesia.

References

- Badan Pusat Statistik. (2021). Retrieved December 29, 2022, from <https://www.bps.go.id/>
- Cohen-addad, V., Kanade, V., Mallmann-trenn, F., & Mathieu, C. (2019). Hierarchical Clustering: Objective Functions and Algorithms. *Journal of the ACM*, 66(4), 26:1-26:42. <https://doi.org/10.1145/3321386>
- Dutta, A. K., Elhoseny, M., Dahiya, V., & Shankar, K. (2020). An efficient hierarchical clustering protocol for multihop Internet of vehicles communication. *Transactions on Emerging Telecommunications Technologies*, 31(5), e3690. <https://doi.org/10.1002/ett.3690>
- Govender, P., & Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Pollution Research*, 11(1), 40–56. <https://doi.org/10.1016/j.apr.2019.09.009>
- Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2018). Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends in Food Science & Technology*, 72, 83–90. <https://doi.org/10.1016/j.tifs.2017.12.006>
- Hashim, H. (2018). Application of Technology in the Digital Era Education. *Research in Counseling and Education*, 02(01). <http://ppsftp.ppi.unp.ac.id/index.php/ijrice/article/view/2/81>
- Oliveira, T. R. A. de, Gravina, G. de A., Oliveira, G. H. F. de, Araújo, K. C., Araújo, L. C. de, Daher, R. F., Vivas, M., Gravina, L. M., & Cruz, D. P. da. (2018). The GT biplot analysis of green bean traits. *Ciência Rural*, 48. <https://doi.org/10.1590/0103-8478cr20170757>
- Roux, M. (2018). A Comparative Study of Divisive and Agglomerative Hierarchical Clustering Algorithms. *Journal of Classification*, 35(2), 345–366. <https://doi.org/10.1007/s00357-018-9259-9>
- Sandi, R. T. (2020). Perkembangan Society 1.0 Hingga Society 5.0. *School of Information Systems*. <https://sis.binus.ac.id/2020/06/09/perkembangan-society-1-0-hingga-society-5-0/>
- Shakeel, P. M., Baskar, S., Dhulipala, V. R. S., & Jaber, M. M. (2018). Cloud based framework for diagnosis of diabetes mellitus using K-means clustering. *Health Information Science and Systems*, 6(1), 16. <https://doi.org/10.1007/s13755-018-0054-0>
- Zeng, K., Ning, M., Wang, Y., & Guo, Y. (2020). *Hierarchical Clustering With Hard-Batch Triplet Loss for Person Re-Identification*. 13657–13665. https://openaccess.thecvf.com/content_CVPR_2020/html/Zeng_Hierarchical_Clustering_With_Hard-Batch_Triplet_Loss_for_Person_Re-Identification_CVPR_2020_paper.html