Hierarchical And Non-Hierarchical Algorithm Clustering: 21st Century Skill Grouping In State Vocational Schools Students

Rahadyan Tajuddien [1] ; Wahyu Nurul Faroh [2]

[1,2] *Universitas Pamulang, Jalan Suryakencana No 1, Tangerang Selatan and 15417, Indonesia*

E-mail: [a] dosen01633@unpam.ac.id
[b] dosen01061@unpam.ac.id

Abstract: Vocational High Schools (SMK), which initially aimed to produce ready-to-work graduates, is the biggest contributor to unemployment. Efforts to overcome this, SMK graduates need to be given 21st century skills to increase agility to survive in the current industrial 4.0 era. 21st century skills are described as: life and career skills, learning and innovation skills and information media and technology skills.

This study conducted a sample grouping of the data from 10,737 students of State Vocational High Schools in Bogor Regency who took advantage of the data mining process using clustering techniques. The method in this study uses the Cross-School Standard Process for Data Mining (CSISP-DM). While the algorithms used for cluster determination are hierarchical and non-heirarchic (K-Means) algorithms. Hierarchical algorithm is a method that does not determine the number of clusters, so that the result or output of a dendogram with a certain number of clusters is determined by the distance until only 2 clusters are formed. K-Means is a non-hierarchical clustering data method that can group student data into several clusters based on the similarity of these data, so that student data with the same characteristics are grouped into one cluster and those with different characteristics are grouped into other clusters. SPSS 25 is used to help find accurate values. The attributes used are life and career skills, learning and innovation skills and information media and technology skills.

The results showed that life and career skills had a frequency of 4,130 students or 38%, learning and innovation skills had a frequency of 4,977 students or 46% and information media and technology skills had a frequency of 3,796 students or 35%. Meanwhile, 21st century skills at 11 SMKN in Bogor Regency are described in stages. The results of Non-Hierarchical Clustering are described in three ways. Furthermore, this research can be used as a basis for decision making to determine future Vocational School development strategies.

Keywords: Clustering, Hierarchical Algorithm, Non-Hierarchical K-Means Algorithm, 21st Century Skills, SMKN

INTRODUCTION

Since Schwab (2016) introduced the term the fourth Industrial Revolution, the transformation of this world has become increasingly rapid. In the past, when the first industrial revolution appeared, the production machines were driven only by the power of water and steam. Furthermore, in the era of the second industrial revolution, the use of electricity for mass production became very massive. Entering the era of the third industrial revolution, the use of electronics and information technology is not something to be denied. So that many productions apply machine automation. In the current era, the fourth industrial revolution, namely the combination of technology by blurring the lines between the physical, digital, and biological fields. The peculiarities of this fourth industrial revolution are the speed, scope, and impact of the system. Because of the growing speed, it has caused disrupting in almost all countries.

Referring to these increasingly massive changes, the education sector certainly needs to move forward with it. Education itself is a basic human foundation to be able to form individuals who are ready to fill and mobilize industrial/business potential in a country. To accommodate this, Education 4.0 was created. Hussin (2018) explains that Education 4.0 is a form of response to the needs of the fourth Industrial Revolution, which aligns humans with technology to create new possibilities.

The urge to fulfill skills through Education 4.0 is something immediately to be realized. But this is not easy. This is even more so if it is implemented simultaneously in the country due to the uneven aspects of ICT (Information Communication and Technology). Muawiyah et al. (2018) explained that ICT itself is a skill that students must have. The implementation of ICT in the learning process will improve the basic ICT skills of students, especially vocational students. Whereas VHS itself, according to Lopez-Mayan and Nicodemo (2012), is a vocational school that is oriented to provide special qualifications to access the labor market.

Data per February 2020 showed that the highest unemployment rate occupied by VHS graduates at 8.49% (Central Statistics Agency, 2020). Every year, VHS produces about 1.4 million graduates, but not all of them enter college or are absorbed in the world of work. This is ironic because VHS, which initially aimed to produce ready-to-work graduates, is the biggest contributor to unemployment. The amount of 11.41% is the largest number of unemployed coming from VHS (Ratnasari, 2018). The same thing was expressed by the Minister of Education and Culture 2016-2019, Muhadjir Effendy, that the quality of VHS in Indonesia is still lacking. As a result, many VHS graduates are unemployed. He even admits that there are only 219 VHSs that have corporation standards (Harususilo, 2018).

The existence of the Covid-19 pandemic has an impact on the decline in economic growth. Of course, this condition greatly affects VHS graduates this year. Many graduates who try to find work are predicted to end up unemployed. Therefore, Ball et al. (2016) directs schools to complement academic skills, and also life and career skills for the success of students when they leave school. Furthermore, the importance of students' mastery of non-cognitive skills, especially as this area is an emerging priority in educational reform and global markets. These are all the scope of 21st-century skills.

Based on the explanation above, researchers need to conduct an assessment of the 21st-century skills possessed by VHS students. Considering that the government has issued a VHS revitalization policy, many VHSs are required to be ready to face the demands of the fourth Industrial Revolution. Then it remains how the VHS can take advantage of these opportunities. Starting here, this research needs to be done.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Twenty First Century Human Resource Skills

Life and Career Skills

Singh & Gera (2015) stated that life skills can be described as a person's ability to behave adaptively and positively that allows him/her to survive by handling life's demands and challenges effectively in his daily life. With these abilities, these individuals can make informed decisions, solve problems, think critically and creatively, communicate effectively, build relationships, empathize with others, manage and cope with their lives in a healthy and productive way.

Life skills themselves are generic, depending on how much life experience a person has lived. However, in the context of vocational students, who are adolescents, their life experiences (challenges) are not as many and complex as adults. It is for this that it is necessary to be honed and improved sequentially. Including when they choose to have a career. These skills will automatically

form a desire in students when completing their formal education. Will he choose to work, continue his studies, or be entrepreneurial? Life and career skills are very pointful in understanding her psychological aspects and maturity. So how the school can provide experience and skill maturation of their students.

Learning and Innovation Skills

As long as humans live their lives, the learning process will always take place. So that humans will always develop and have a lot of thoughts. This is what makes learning and innovation skills always emerge. Paiwithayasiritham (2014) explains that learning and innovation skills are divided into creativity and innovation, critical thinking, and solving problems, and communication and collaboration. For this reason, learning and innovation skills are included as skills in the 21st century. Often in schools, these learning and innovation skills are supported by qualified school facilities, and what about schools with minimal facilities. This is what needs to be examined through this research.

Information Communication and Technology (ICT)

Current technological developments require changes in teaching methods and student perspectives. ICT will certainly be implemented and run by many schools. Even these ICT facilities are always highlighted as the excellence of the school. But not necessarily these advantages can be enjoyed and understood by all students. On the other hand, it is imperative to be part of the skills of the 21st century. Today, technology cannot be separated from each other in their daily lives. So, technology is a part of a student's life and is a must to be able to master it. Many cases of ICT can make it easier for students to learn. However, so far the ICT is controlled by VHS students. For this reason, it is necessary to explore the meaning of ICT in these students. In this research, ICT can be described as learning that applies the concepts of computer and multimedia learning. The application of this concept will be more interesting, fun, and believed to be able to improve learning outcomes (Nurdyansyah & Riananda, 2016). On this basis, the need for a vocational school student to master ICT.

Data Mining

Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from large databases. According to Larose (2006), data mining is divided into several groups based on the tasks that can be done.

a. Description

Sometimes researchers and analysts simply want to try to find ways to describe the patterns and trends contained in the data.

b. Estimate

Estimation is almost the same as classification, except that the estimated target variable is more numerical than categorical.

c. Prediction

Prediction is almost the same as classification and estimation, except that in the prediction the value of the results will be in the future.

d. Classification

In classification, there are target categorical variables.

e. Clustering

Clustering is a method for finding and grouping data that have similar characteristics (similarity) between one data and another. Clustering is an unsupervised data mining method.

f. Association

The task of associations in data mining is to find attributes that appear at a time. In the business world, it is more commonly called shopping cart analysis.

Clustering

Clustering is a data mining technique that functions to classify data based on the similarity of data in one group and to minimize similarities in other groups. Clustering is a grouping of records, observations, or attention and forms a class of similar objects. A cluster

is a collection of records that have similarities with one another and have similarities with records in other clusters. There are two clustering methods that we know, namely hierarchical clustering and non-hierarchical clustering.

Hierarchical clustering has two types of methods, namely agglomerative and divisive. Agglomerative clustering or also called the "bottom-up" method because each object is considered a single cluster than the clusters are combined so that only one cluster remains. Divisive clustering is also called the "top-down" method. In the divisive method, initially, all objects are considered to be the same cluster, thus the process of splitting the clusters into two clusters is carried out until each object is considered a single cluster. The process of forming clusters in hierarchical clustering is described through a two-dimensional diagram called a dendrogram. The dendrogram image is the formation of clusters with both agglomerative and divisive procedures in the form of a dendrogram. The procedure that is often used in the hierarchical clustering method is the agglomerative procedure. Initially, there are n members/observations which are considered n clusters or single groups and in the end produce one cluster or one group containing n members. Commonly used basic join operations include single linkage, complete linkage, and average linkage. The three methods have differences in determining the distance (similarity) between objects and groups of objects or groups of objects with groups of objects. Single linkage is a merging operation that uses the minimum distance or nearest neighbor. Conversely, complete linkage uses the farthest distance (farthest neighbor). Meanwhile, average linkage uses a distance average to determine the distance between objects. Other combined operating methods include Ward's minimal variance, centroid method, median method, and average linkage weighted.


Hierarchical Dendrogram
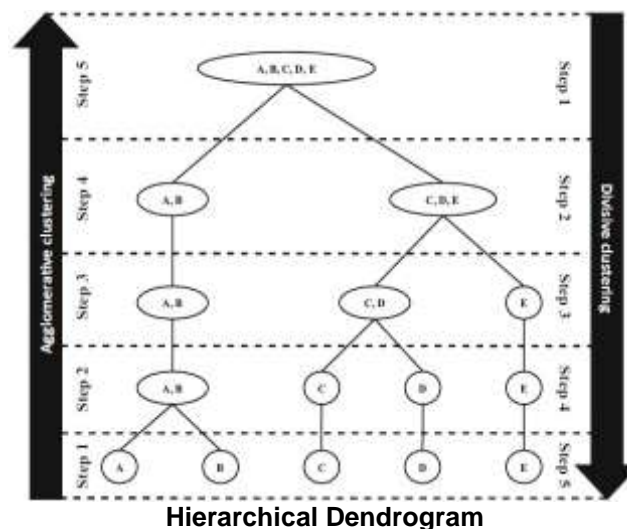
K-means clustering is one of the Non-Hierarchical Clustering data methods that classify data in the form of one or more clusters/groups. Data that have the same characteristics are grouped into one cluster/group, and data that have different characteristics are grouped with other clusters/groups so that data in one cluster/group has a small level of variation. The steps for clustering using the K-Means method are as follows:

1. Select the number of clusters K
2. Iinitialization of the cluster center can be done in various ways. But what is most often done is random. Cluster centers are assigned an initial value with random numbers.
3. Allocate all data/objects to the nearest cluster. The proximity of two objects is determined based on the distance between the two objects. Likewise, the closeness of data to a particular cluster is determined by the distance between the data and the cluster center. In this stage, it is necessary to calculate the distance of each data to each cluster center. The distance between one data and one particular cluster will

determine which data is included in which cluster. To calculate the distance of all data to each cluster center point, the Euclidean distance theory can be used.

4. Recalculate the cluster center with the current cluster membership. The cluster center is the average of all data/objects in a particular cluster. If desired, you can also use the median (middle value) of the cluster. So the mean (mean) is not the only measure that can be used.

5. Assign each object again using the new cluster center. If the cluster center does not change again, the clustering process is complete. Or, go back to step number 3 until the cluster center does not change anymore.

Meanwhile, the framework for this research is reflected as follows:
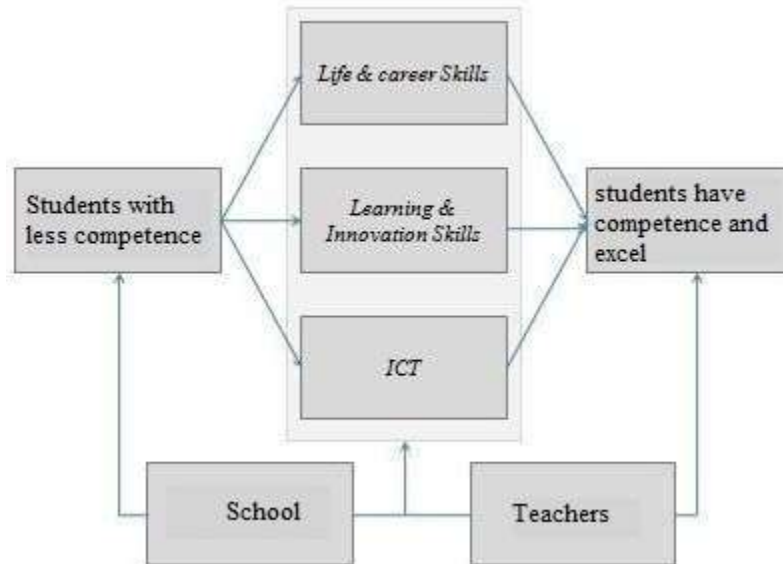
Figure Framework

METHODS

This research uses quantitative data. The data used are real numbers that describe the attributes in this study. While the analysis method used is Modified Hierarchical Clustering and Non-Hierarchical Clustering, where the researcher creates several clusters. These are then made into defined clusters. While the data used is in the form of data mining, which must be divided into groups based on clustering, a method for finding and classifying data.

Larose (2006) then describes the phases in data mining known as CRISP-DM (Cross School Standard Processor Data Mining) as shown in Figure 3.1 are as follow:

a. School Understanding Phase
b. Data Understanding Phase
c. Data processing phase (Data Preparation Phase)
d. Modeling Phase
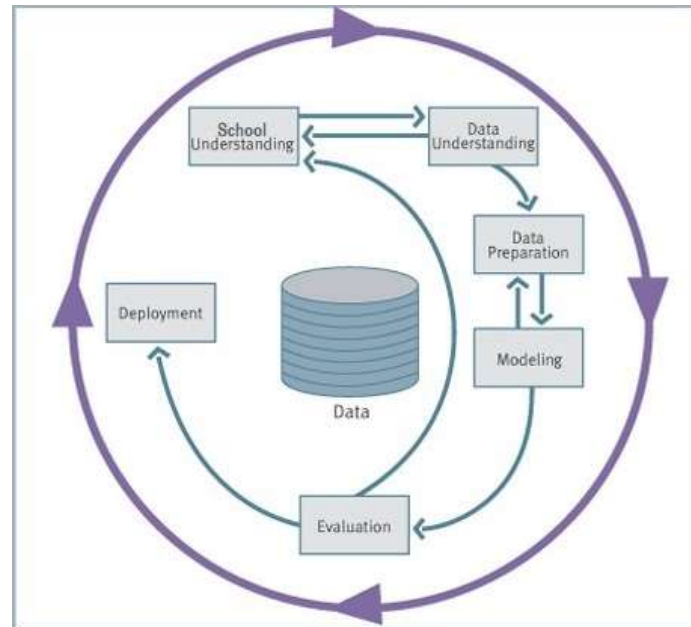e. Evaluation Phase
f. Deployment Phase

Figure CSISP-DM diagram

RESULT AND DISCUSSION

School understanding, the first stage in the CSISP-DM process which can also be called the School understanding stage (research). In this stage, the researcher explores the need for data search and clustering according to school objectives.

a. Determining School Goals

The purpose of the school-based data processing on vocational school students in 2020, among others, is to classify 21st-century human resource skills for students in 2020. Data grouping is carried out because the data in the field shows that the highest number of unemployed is VHS students graduate as much as 8.49% as of February in 2020. The purpose of grouping data can be used as one of the bases for decision making to determine the VHS development strategy.

b. Conduct a Situation Assessment

At this stage, an understanding of the school's goals is needed and translating it into data mining objectives.

c. Determine the initial strategy

The initial strategy in implementing the objective of data mining was to first collect student data from 11 VHSNs in Bogor Regency.

d. Data Understanding

The dataset of VHSN students obtained via google form from 11 VHSN in Bogor Regency are processed into excel document tabulation data containing data on 21st Century HR Skills for student.

Initial data collection, The collection of data sources using google form which is linked to all students on 11 public VHS in Bogor Regency. The data collected is then processed into data used in this study. This is the tabulation dataset for grade 10, 11, and 12 2020 SMKN students with the .xlsx format.

a. Describe data

The VHS student dataset consists of the attributes of life and career skills, learning and innovation skills, and information media and technology skills. The amount of data on the attribute is 10,737 records.

b. Evaluate data quality

The results of the evaluation of data quality were that student respondents who answered not according to competence were not considered in frequency and found several null values called missing values in the attributes in the student datasets.

c. Attribute Selection

The attributes used are:

1. Life and career skills,
    a. Can adapt to the work environment
    b. Can work (career / open a business)
    c. Can handle challenges
    d. Can make decisions
    e. Can solve the problem
2. Learning and innovation skills (learning and innovation skills)
    a. Always interested in learning new things
    b. Innovate often
    c. Creative
    d. Critical thinking
    e. Can cooperate
3. Information media and technology skills (information media and technology skills)
    a. Using ICT learning methods
    b. ICT Facilities in Schools
    c. Everyday using ICT
    d. Can take advantage of ICT
    e. Teachers Have ICT Competence

Data preparation includes all activities to build student datasets that will be applied to modeling tools, from the initial raw data in the form of student datasets, and will then carry out the data mining process.

a. Data Selection

The data used are students' positive answers related to life and career skills, shortened to LNCS, learning and innovation skills shortened to LNIS, and information media and technology skills shortened to IMNTS.

b. Raw data processing (Preprocessing Data)

At this stage, it is the stage to ensure that the selected VHS student data is suitable for data processing.

c. Data transformation

Data of nominal types such as skills names and school names must be initialized in numerical form. Initialization can be done by:

1. In the Skill Name, initially the division is carried out. The skills are divided into Skills sections.

Table shows that it can be concluded that the highest percentage of skills mastered by VHS students is learning and innovation skills (LNIS) with 46% of the data. Meanwhile, the lowest percentage of skills is in information media and technology skills (IMNTS) within a total of 46%.

2. Hence the school names are divided into several parts.

Table shows that the highest percentage of students' skills seen is SMKN 1 Gunungputri 48% and SMKN 1 Bojong Gede 48%. Meanwhile, the data for the lowest percentage of student skills is at SMKN 1 Puncak Cisarua 24%.

Table Skill Name Initialization

Skill Name	Student Samples	Frequency	Percentage	Initialization
Life and career skills (LNCS)	10,737	4,130	38%	1
Learning and innovation skills (LNIS)	10,737	4,977	46%	2
Information media and technology skills	10,737	3,796	35%	3

Table Initialization of School Names

VHS	∑ Students	Freq.	%	Initialization
SMKN 1 Cileungsi	1,117	394	35%	1
SMKN 1 Puncak Cisarua	602	145	24%	2
SMKN 1 Gunung Sindur	888	298	34%	3
SMKN 1 Cariu	893	273	31%	4
SMKN 1 Gunungputri	1,345	645	48%	5
SMKN 1 Ciomas	928	402	43%	6
SMKN 1 Bojong Gede	1,014	488	48%	7
SMKN 1 Kemang	307	85	28%	8
SMKN 1 Leuwiliang	861	260	30%	9
SMKN 1 Cibinong	2,022	810	40%	10
SMKN 2 Cibinong	760	330	43%	11

1. Making variables and types of school names and skills names.

Table shows the number of 21st-century skills per attribute and school. The data shows that the average number of students with the highest number of students who have 21st-century skills is SMKN 1 Cibinong with a total of 810 students. Furthermore, in the second place, there is SMKN 1 Gunungputri with a total of 645 students. Meanwhile, the school that has several 21st-century skilled students is SMKN 1 Kemang with a total of 85 students.

Table Skilled 21st Century Students

VHS	21st Century Skilled Students			
	Life And Career Skills (LNCS)	Learning And Innovation Skills (LNIS)	Information Media And Technology Skills (IMNTS)	Mean
SMKN 1 Cileungsi	363	524	297	394
SMKN 1 Puncak Cisarua	145	138	152	145
SMKN 1 Gunung Sindur	240	361	293	298
SMKN 1 Cariu	248	320	250	273
SMKN 1 Gunungputri	609	757	570	645
SMKN 1 Ciomas	489	506	402	466
SMKN 1 Bojong Gede	669	625	488	594
SMKN 1 Kemang	81	102	73	85
SMKN 1 Leuwiliang	340	220	220	260
SMKN 1 Cibinong	686	1028	717	810
SMKN 2 Cibinong	260	397	333	330

2. The statistical descriptive output is used to show the minimum, maximum, and mean data. The data shows that the Standard Deviation value is away from the mean, which means that the data distribution is varied.

Table Descriptive Statistics

	N	Min.	Max.	Mean	Std. Deviation
LNCS	11	81	686	375.45	209.569
LNIS	11	102	1,028	452.55	276.853

| IMNTS | 11 | 73 | 717 | 345.00 | 188.080 |

Table shows the minimum and maximum frequency of each skill of all VHS in the Bogor District. There is a minimum number of Information Media and Technology Skills with a total of 73 students.

Modeling, a phase that directly involves data mining techniques, by selecting data mining techniques and determining the algorithms to be used. Selection of the modeling technique Tool used is SPSS version 25:
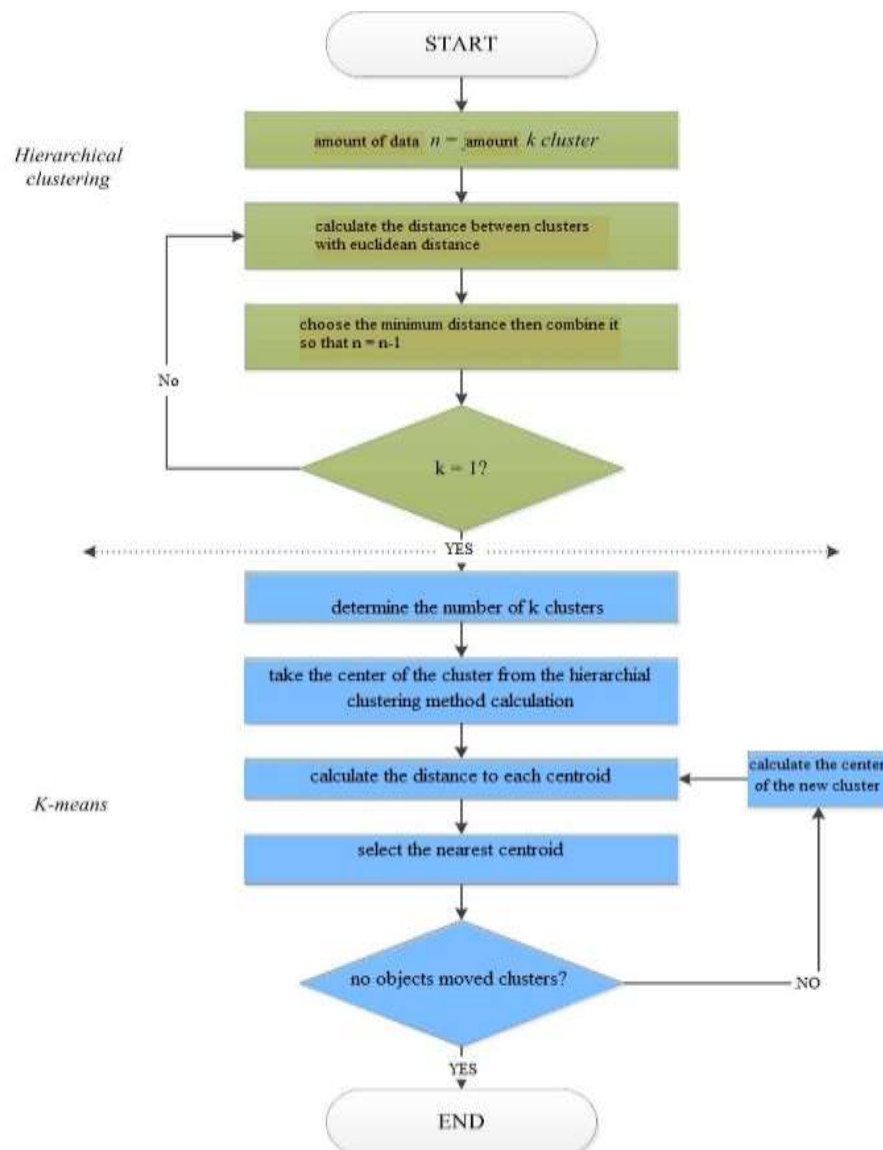
Figure Flowcart SPSS

Figure shows the steps in data processing.

a. Hierarchical Clustering

The results in Figure 4.2 are the dendrogram results of cluster analysis with the single linkage method, where the more clusters are selected, the smaller the distance. From the dendrogram with a distance of more than 25, 1 cluster is obtained (no grouping occurs), whereas if the distance is 25 then 2 clusters are obtained, namely cluster 1 consisting of

SMKN 1 Cibinong while other schools enter into cluster 2. This result is similar to the previous output table (cluster membership table) if 2 clusters are selected..

Figure Dendrogram

Likewise, if we take the distance of 20, there are 3 clusters, where cluster 1 consists of SMKN 1 Cileungsi, SMKN 1 Puncak Cisarua, SMKN 1 Gunung Sindur, SMKN 1 Cariu; cluster 2 consists of SMKN 1 Gunungputri and SMKN 1 Bojong Gede; while the remaining SMKN are included in cluster 3. Equally, for other distances, which result in the number of clusters that may differ from one another. The more clusters formed, the smaller the distance. That is, the more similar the clusters are and have the similarity between clusters.

In processing cluster analysis using SPSS, it can be used as a note that the dendrogram results in the hierarchical method are not determined by the number of clusters. The result or dendrogram output with a certain number of clusters is determined by the distance until only 2 clusters are formed. In the dendrogram (figure 4.2), the maximum distance is 25 (i.e. until only 2 clusters are formed). In contrast to non-hierarchical methods such as K-Means, the number of clusters is predetermined..

b. Non-Hierarchical Clustering

The Initial Clusters table shows the first three clusters formed. Then the K-Means Cluster method will test and reallocate the existing clusters. This process is called Iteration, which contains changes to the initial cluster in table.

Table Initial Clusters

Initial Cluster Centers

	Cluster 1	Cluster 2	Cluster 3
Zscore(LNCS)	1.48183	-1.40505	-.55091
Zscore(LNIS)	2.07856	-1.26618	-.20063
Zscore(IMNTS)	1.97788	-1.44619	-.06380

Source: data processing SPSS 25

Table Iteration History

Iteration History[a]

	Change in Cluster Centers		
Iteration	1	2	3
1	1.061	.268	.353
2	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 1.943.

Source: data processing SPSS 25

The result of K-Means is the Final Cluster Centers, which contain three clusters to divide 11 Public VHSs based on the life and career skills shortened to LNCS, learning and innovation skills shortened to LNIS and media skills information and technology (information media and technology skills) will be abbreviated as IMNTS in 2020.

Table Final Clusters

Final Cluster Centers

	Cluster		
	1	2	3
Zscore(LNCS)	1.33232	-1.25236	-.24871
Zscore(LNIS)	1.26705	-1.20116	-.23314
Zscore(IMNTS)	1.31150	-1.23617	-.24369

Source: data processing SPSS 25

Based on the final cluster center table, it can be concluded that of the three variables that are relevant to differentiating the contents of clusters in grouping Public VHSs which include cluster 1, cluster 2, and cluster 3. The data in Table shows that cluster 1 contains 3 VHSs, while cluster 2 contains 2 VHSs and cluster 3 contains 6 VHSs.

Table Number of Cases

Number of Cases in each Cluster

Cluster	1	3.000
	2	2.000
	3	6.000
Valid		11.000
Missing		.000

Source: data processing SPSS 25

Evaluation is an advanced phase of data mining objectives. The evaluation is carried out in-depth with the aim that the results at the modeling stage match the goals to be achieved in the school understanding stage.

a. Evaluation Results

This stage assesses the extent to which the results of data mining modeling meet the data mining objectives that have been determined at the School understanding stage.

b. Review Process

At this stage, the researcher ensures that all-important stages/factors that have been carried out in data processing are not overlooked.

c. Determine Next Steps

At this stage is the stage in determining the next steps to be taken. There are 2 options, returning to the initial stage (School understanding) or continuing to the final stage (deployment).

Deployment, this research eventually can see the distribution of the data that has been processed into clusters.

a. Hierarchical Clustering Analysis

Details of the number of member clusters formed can be seen in table:

Table Hierarchical Clustering

Case	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1:SMKN 1 Cileungsi	1	1	1	1
2:SMKN 1 Puncak Cisarua	2	2	1	1
3:SMKN 1 Gunung Sindur	1	1	1	1
4:SMKN 1 Cariu	1	1	1	1
5:SMKN 1 Gunungputri	3	3	2	1
6:SMKN 1 Ciomas	4	1	1	1
7:SMKN 1 Bojong Gede	3	3	2	1
8:SMKN 1 Kemang	2	2	1	1
9:SMKN 1 Leuwiliang	1	1	1	1
10:SMKN 1 Cibinong	5	4	3	2
11:SMKN 2 Cibinong	1	1	1	1

Source: data processing

The output results show the grouping of the objects of observation against the 5 clusters that have been preset. There is no specific reason why choosing 5 clusters. This is done so that the results are more centered on 2, 3, 4, or 5 clusters so that they are easy to read. For example, if 2 clusters are used, then SMKN 1 Cileungsi, SMKN 1 Puncak Cisarua, SMKN 1 Gunung Sindur, SMKN 1 Cariu, SMKN 1 Gunungputri, SMKN 1 Ciomas, SMKN 1 Bojong Gede, SMKN 1 Kemang, SMKN 1 Leuwiliang and SMKN 2 Cibinong are at cluster 1 while the other SMKNs are grouped into cluster 2. Likewise for the number of clusters as many as 3, 4, or 5.

For 5 clusters, the following groupings were obtained: Cluster 1 consisted of SMKN 1 Cileungsi, SMKN 1 Gunung Sindur, SMKN 1 Cariu, SMKN 1 Leuwiliang and SMKN 2 Cibinong. Cluster 2 consists of SMKN 1 Puncak Cisarua and SMKN 1 Kemang. Cluster 3 consists of SMKN 1 Gunungputri and SMKN 1 Bojong Gede. Cluster 4 consists of SMKN 1 Ciomas. Finally, Cluster 5 is only SMKN 1 Cibinong.

b. Non-Hierarchical Clustering Analysis

As previously explained, the non-hierarchical method is a clustering method where the number of clusters is known. In this case, the number of clusters is determined to be 3 clusters. From 4.10 it can be seen that the number of objects in cluster 1 (first) is 3 SMKN in Bogor Regency while cluster 2 (second) is 2 SMKN in Bogor Regency and the rest are grouped into cluster 3 (third). This can be described as follows:

Cluster 1: SMKN 1 Gunungputri, SMKN 1 Bojong Gede and SMKN 1 Cibinong;

Cluster 2: SMKN 1 Puncak Cisarua and SMKN 1 Kemang;

Cluster 3: SMKN 1 Cileungsi, SMKN 1 Gunung Sindur, SMKN 1 Cariu, SMKN 1 Ciomas, SMKN 1 Leuwiliang and SMKN 2 Cibinong

CONCLUSIONS

The results of this study can be concluded that the percentage of 21st-century skills with a sample of 10,737 students in public VHSs throughout Bogor Regency is life and career skills with a frequency of 4,130 students or 38%; learning and innovation skills with a frequency of 4,977 students or 46%; and information media and technology skills with a frequency of 3,796 students or 35%. This supports the picture by Lawitta et al. (2017), where there is still a need to improve the skills of VHS students in Indonesia. Moreover, not all VHS students approach the expected 21st-century skills.

Meanwhile, 21st-century skills in students at each public VHSs in Bogor Regency are respectively SMKN 1 Gunungputri and SMKN 1 Bojong Gede by 48%; SMKN 2 Cibinong and SMKN 1 Ciomas respectively by 43%; SMKN 1 Cibinong by 40%; SMKN 1 Cileungsi by 35%; SMKN 1 Gunung Sindur by 34%; SMKN 1 Cariu by 31%; SMKN 1 Leuwiliang by 30%; SMKN 1 Kemang is 28% and the last is SMKN 1 Puncak Cisarua by 24%. These results indicate, not a single public VHS in Bogor Regency has been able to approach the expected 21st-century capabilities.

Felestin and Triyono (2015) explain that the implementation of TQM in public VHS is better than private VHS in Indonesia. This is because of the funding factor from the government. However, when collided with the results of the above research, it shows that school management is not optimal yet. Several aspects that need to be addressed include teacher competence, especially in ICT (Nurhabibah et al., 2017) and professionalism (Agung, 2018), equipment (Suartini, 2019), and vocational learning (Ermawati and Wagiran, 2018). Hopefully, there are no problems with other aspects.

The results of non-hierarchical clustering are that in Cluster 1 there are schools SMKN 1 Gunungputri, SMKN 1 Bojong Gede and SMKN 1 Cibinong. Meanwhile, Cluster 2 of SMKN 1 Puncak Cisarua and SMKN 1 Kemang. Cluster 3 SMKN 1 Cileungsi, SMKN 1 Gunung Sindur, SMKN 1 Cariu, SMKN 1 Ciomas, SMKN 1 Leuwiliang and SMKN 2 Cibinong.

Although it only uses data from public VHSs in all Bogor districts, the results of this study are a representation of the condition of public VHS in Indonesia. So that in the end, the results of this study can be used as a basis for decision making to determine the strategy for developing vocational schools in Indonesia.

VHS as a formal educational institution that is expected to be able to support the acceleration of national development must be sensitive to its potential. Based on basic education data (Dapodik), the total number of VHSs is currently 14 thousand units, 25% of which are public VHSs and the rest are private VHSs. While the number of students is more than 5 million students, with 9 areas of expertise, 49 expertise programs, and 146 skill competencies. In addition to being an indicator of success in providing access to secondary education, these figures also reflect the magnitude of the challenges for the government to ensure the provision of quality education following national education standards to overcome industrial and work problems for VHS graduates as follows:

1. There is a gap between expectations and reality between high school graduates and the needs of industry and the world of work
2. There is no link and match between graduates and needs
3. Hard skills and soft skills are not yet balanced.
4. Lack of industrial training materials for VHS graduates.

The solution, researchers divided their respective roles as follows:

1. Government have to run the Vocational Working Committee Program, the White Paper Program, and creating policies and competency ecosystem for VHS graduates.
2. VHS needs to compile the curriculum together with Campuses and Industries, taught by industry practitioners, run the apprenticeship program, certificate teachers and students competencies certification through training, assistive equipment, and stimulate research student competencies.
3. Industrial preferably to recruit the vocational graduate, conduct apprenticeship program, provide scholarship program and official ties, and provide research for vocational sectors.

ACKNOWLEDGEMENT

REFERENCE

Agung, I. (2018). Improvement of Teacher Competence and Professionalism and School Management Development in Indonesia. American Journal of Educational Research, 6(10), 1388-1396. https://doi.org/10.12691/education-6-10-8

Ariyanto, A., Nuryani, A., & Sunarsi, D. (2020). Pengaruh Store Atmosphere Dan Promosi Terhadap Keputusan Pembelian Pada Alfamart BSD Tangerang Selatan. Jurnal Ekonomi Efektif, 3(1).

Ball, A., Joyce, H., and Anderson-Butcher, D. (2016). Exploring 21st Century Skills and Learning Environments for Middle School Youth. International Journal of School Social Work, 1(1), 1-25. https://doi.org/10.4148/2161-4148.1012

Effendy, A., & Sunarsi, D. (2020). Persepsi Mahasiswa Terhadap Kemampuan Dalam Mendirikan UMKM Dan Efektivitas Promosi Melalui Online Di Kota Tangerang Selatan. Jurnal Ilmiah MEA (Manajemen, Ekonomi, & Akuntansi), 4(3), 702-714. https://doi.org/10.31955/mea.vol4.iss3.pp702-714

Ermawati, R. and Wagiran (2018). Profile of Vocational Learning in the Era of Industrial Revolution 4.0 (Studies at Department of Automotive Vocational High School). Journal of Physics: Conference Series: Vol. 1273 (pp. 1-10). International Conference on Vocational Education of Mechanical and Automotive Technology. https:// doi:10.1088/1742-6596/1273/1/012015

Felestin and Triyono, M. (2015). The Implementation of Total Quality Management at Vocational High Schools in Indonesia. Research and Evaluation in Education Journal, 1(1), 13-24. https://doi.org/10.21831/reid.v1i1.4895

Haque, MG., Munawaroh, Sunarsi, D., (2020). Analysis of SMEs Culinary Marketing Strategy During Covid 19 Pancemic: A Study at "Sate Bebek Cilegon" Resto in Cilegon, Banten. International Journal of Education, Information Technology, and Others. Vol.3. Issue 2

Harususilo, Y. (2018, April 26). Ini Alasan Lulusan SMK Banyak Menganggur. Kompas.com. https://edukasi.kompas.com/read/2018/04/26/15342831/ini-alasan-lulusan-smk-banyak-menganggur (Retrieved on February 7th, 2019)

Hussin, A. (2018). Education 4.0 Made Simple: Ideas for Teaching. International Journal of Education & Literacy Studies, 6 (3), 92-98. https://doi.org/10.7575/aiac.ijels.v.6n.3p.92

Kasmad, K., Mustakim, M., & Sunarsi, D. (2020). Increasing Community School Interest Through Service Quality, Prices and Promotion in Vocational High Schools. Journal of Educational Science and Technology (EST), 6(2).

Larose, D. (2006). Data Mining Methods and Models. New Jersey: John Wiley & Sons, Inc.

Lawitta, R., Sihaloho, L. and Arianti, J. (2017). Vocational High School in Indonesia Facing ASEAN Economic Community (AEC). Advanced in Social Science, Education and Humanities Reseacrh (ASSEHR): Vol. 158 (pp. 950-957). International Conference on Teacher Training and Education. https://doi.org/10.2991/ictte-17.2017.28

Lopez-Mayan, C. and Nicodemo, C. (2012, January). Vocational High School or Vocational College? Comparing the Transitions from School to Work [Paper presentation]. IZA Discussion Paper No 6309, Bonn, Germany

Lukiastuti, Fitri, et.al (2020). The Influence of Entrepreneur's Personal Characteristics on SMES Performance Mediated by Entrepreneurial Orientation. International Journal of Psychosocial Rehabilitation. Volume 24 - Issue 8

Maddinsyah, A., Sunarsi, D., Hermawati, R., Pranoto. (2020). Analysis of location selection effect on the user decision that influcence the success of the service business of micro, small and

medium enterprise (MSME) in bandung timur region. International Journal of Advanced Science and Technology. Vol. 29 No. 06

Muawiyah, D., Yamtinah, S. dan Indriyanti N. (2018). Higher Education 4.0: Assessment on Environtmental Chemistry Course in Blended Learning Design. IOP Conference Series: Journal of Physics: Conf. Vol. 1097 (pp. 1-7).International Conference on Research, Implementation, & Education of Mathematics and Sciences. https://doi :10.1088/1742-6596/1097/1/012058

Nurdyansyah and Riananda, L. (2016). Developing ICT-Based Learning Model to Improve Learning Outcomes IPA of SD Fish Market in Sidoarjo. Jurnal Teknologi Pendidikan: Vol. 1 (pp. 929-940). Proceedings of International Research Clinic & Scientific Publications of Educational Technology. https://journal.unesa.ac.id/index.php/jtp/article/view/1137/781

Nurhabibah, Setiawan, A., Yanti, H., Miraj, Y., dan Yannuar (2017). Analysis of ICT Literacy Competence among Vocational High School Teachers. IOP Conferences Series: Materials Sciences and Engineering: Vol. 306 (pp. 1-7). International Conference on Innovation in Engineering and Vocational Education. https://doi:10.1088/1757-899X/306/1/012097

Paiwithayasiritham, C. (2014). Factor Analysis of The 21st Century Learning and Innovation Skills of The Teaching Professional Students. Veridian E-Journal, 7(5), 27-35. http://www.educ.su.ac.th/images/journal/web/57-1/11.pdf

Purwanti, P., Sarwani, S., & Sunarsi, D. (2020). Pengaruh Inovasi Produk Dan Brand Awareness Terhadap Keputusan Pembelian Konsumen Pada PT. Unilever Indonesia. Inovasi, 7(1), 24-31.

Purwanto, H., Fauzi, . M., Wijayanti, . R., Awwaly, . K. U. A., Jayanto, . I., Mahyuddin, ., Purwanto, . A., Fahlevi, . M., Adinugraha, H. H., Syamsudin, . R. A., Pratama, . A., Ariyanto, . N., Sunarsi, . D., Hartuti, . E. T. K. & Jasmani, . (2020) Developing Model of Halal Food Purchase Intention among Indonesian Non-Muslim Consumers: An Explanatory Sequential Mixed Methods Research. Systematic Reviews in Pharmacy, 11 (10), 396-407. doi:10.31838/srp.2020.10.63

Ratnasari, E. (2018, February 26). BPS: Lulusan SMK Banyak Menganggur Sepanjang 2017. cnnindonesia.com. https://www.cnnindonesia.com/gaya-hidup/20180223141505-282-278333/bps-lulusan-smk-banyak-menganggur-sepanjang-2017 (Retrieved February 7th, 2019).

Schwab, K. (2016, Januari 14). The Fourth Industrial Revolution: What It Means and How to Respond. Weforum.org. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/ (Retrieved on March 3rd, 2019).

Singh, H. & Gera, M. (2015). Strategies for Development of Life Skills and global Competencies. International Journal of Scientific Research, 4(6), 760-763. https://www.doi.org/10.36106/ijsr

Suartini, T. (2019). Influence Application of Learning Model on Vocational Education Based on Quality Issurance. SAGE Open, 9(2), 1-11. https://doi.org/10.1177/2158244019851552

Sunarsi, D., & Baharuddin, A. (2019). The Effect of Service Quality and Price Accuracy on Consumer Confidence and Implications for Sales Increase. PINISI Discretion Review, 3(2), 101-110.