Clustering Methods Based on Indicator Process Model to Identify Indonesian Class Hospital

Linda Septiana Hariyanti
Departement of Business Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
lindaseptianah98@gmail.com

Kelly Rossa Sungkono
Departement of Informatics Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
kelly@its.ac.id

Riyanarto Sarno
Departement of Informatics Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
riyanarto@if.its.ac.id

***Abstract*— Health is the most important aspect of human life. One institution that provides health services for individuals who fully understand the road, hospitalization, and emergency care hospitals. According to Law No. 44 of 2009 in Indonesia, the hospital becomes 4 classes, namely class a, class b, class c, and class d. The indicators for each hospital have been determined. This indicator is human resources such as doctors, pharmacists, and time. The study was conducted to determine the value delivered by the hospital per indicator through a direct hospital process model. First, log events directly from the hospital process based on hospital service indicators. The results of this process will get the final value of each indicator that has been determined. These values are then analyzed using the k-means cluster method, hierarchical complete linkage, and hierarchical single linkage to get classes from the hospital. Based on the analysis that has been done, it can be seen that the k-means cluster method has better accuracy, precision, and recall that a hierarchical complete linkage, and hierarchical single linkage.**

Keywords—hierarchical complete linkage, hierarchical single linkage, k-means cluster, process model.

I. INTRODUCTION

Health is the most important aspect of human life. Every activity to improve health and the place used is called a health facility. Hospitals are one of the health facilities that provide curative and preventive health services for the community, where services provided are inpatient, outpatient, and emergency care [1]. Hospital classes in Indonesia are classified into 4, namely class a hospital, class b hospital, class c hospital, and class d hospital. The differences from each class are based on service, employment, physical and equipment elements with each class hospitals that have predetermined indicators. One of the indicators used in this study is the indicator of human resources and time criteria. Indicators of human resources used include doctors and pharmacists in outpatient, inpatient, and emergency units. While the time criteria used include waiting time in outpatient and waiting time to take prescription drugs.

Indicator values can be obtained from the hospital process model through direct simulations that have been carried out. The process of this model is used, identifying possible anomalies, monitors and control operations [2]. In this study using the hospital process model. The process model itself has been widely used to measure company process performance and show problems related to the process. Process models can be drawn manually by analysts or can be generated based on event logs [3]. Event logs are used to analyze problems in business processes [4]. The process model also plays an important role in facilitating communication between various stakeholders and in documenting organizational business processes [5]. This business process has information about where and when activities are carried out, input and output activities, initial conditions before activities are carried out and final conditions after the activity is executed [6].

To get a hospital class, researchers used the k-means clustering method, complete linkage hierarchy, and hierarchical single linkage. The k-means cluster analysis of N groups of data points into the cluster k by minimizing the number of squared distances between each point and the nearest cluster average [7]. The initial grouping in k-means starts with a large cluster with one center then divided into two clusters, which means that the new cluster is still learning iteratively. This process continues until the specified cluster is obtained [8]. Single linkage is one example of a form of data grouping. Hierarchical single linkage is formed with conduct discussions without determining the number of groups first and groups forming groups with short groups [9]. This grouping method uses the object that is closest almost equal between objects with each other to be grouped. Distance between two clusters was chosen from the closest distance from all data pairs in the two clusters [10]. Hierarchical complete linkage method is a clustering method similar to the single linkage clustering method but uses the opposite approach [10]. The hierarchical complete linkage method is formed with doing an approach without first determining the number of groups and determining distances are in the longest distance (smallest similarity) with each other [9].

The data used is simulation data which amounts to 40 data. The data were analyzed to find accuracy, precision, and recall using two methods, namely complete linkage hierarchy, single linkage hierarchy and k-means cluster, which is known which method provides the most appropriate accuracy, precision, and recall where the method will be used to classify existing classes in the hospital. Precision is the level of accuracy between user requests and answers system. Accuration is a comparison between information answered by the system correctly with all information. A recall is a measure of accuracy between the same information as information had been called before [11].

II. LITERATURE REVIEW

The methods used in analyzing the types of hospitals in Indonesia are as follows.

A. Hospital Classification

Based on the Law concerning hospitals no.44 in 2009, the hospital was a health service institution that held individual

978-1-7281-3832-9/19/$31.00 ©2019 IEEE

health services that provided inpatient, outpatient, and emergency services. Hospital classification is a class of hospital grouping based on facilities and service capabilities [2]. Hospital classification based on facilities and service capabilities, general hospitals as follows.

1) Class A general hospitals are general hospitals that have extensive facilities and broad medical and subspecialists medical service capabilities.
2) Class B general hospitals are public hospitals that have medical service facilities and capabilities of at least limited specialists and subspecialists.
3) Class C general hospital is a public hospital that has the facilities and capabilities of basic specialist medical services.
4) General hospital class D general hospitals that have basic medical facilities and capabilities.

B. K-Means Cluster

Non-hierarchy methods (k-means cluster) begin by selecting a number of initial cluster values that correspond to the desired number and then the objects are combined into these clusters. Grouping using the k-means method is based on the value of its membership function. The membership function is based on the minimum distance between objects with the cluster center (centroid) [1].

The k-means algorithm aims to minimize objective functions which are error squared functions. For example, there are n objects and p variables. The distance between the object and the group is calculated using the Euclidean squared distance is described in Eq. (1)

$$d_{(i,k)} = \sum_{k=1}^{n} [x_{(i,k)} - \bar{x}_{(l,k)}]^2 \quad (1)$$

with :

$i = 1,2,....,n$

$k = 1,2,....,m$

C. Hierarchical Clustering

Hierarchical clustering is a recursive partitioning of a dataset into successively smaller clusters. The input is a weighted graph whose edge weights represent pairwise similarities or dissimilarities between data points. Hierarchical clustering is represented by a rooted tree where each leaf represents a data point and each internal node represents a cluster containing its descendant leaves. Computing a hierarchical clustering is a fundamental problem in data analysis, it is routinely used to analyze, classify, and pre-process large datasets [12]. Following is the hierarchy grouping method used.

1) Hierarchical Complete Linkage

Complete linkage provides assurance that all items in one cluster are at the farthest distance (the smallest similarity) from each other. Agglomerative algorithms generally begin by specifying entries (matrix elements) in $D = \{dik\}$ and combining corresponding objects such as U and V to get clusters (UV) [2]. The steps of the algorithm above the distances between clusters (UV) and other W clusters can be show in Eq. (2)

$$d_{(UV)W} = \mathbf{min}[d_{UW}, d_{VW}] \quad (2)$$

Information :

d_{uv} = the distance between closest neighbors of clusters U and W

d_{vw} = the distance between the closest neighbors of clusters V and W

2) Hierarchical Single Linkage

Inputs for single linkage algorithms can be in the form of distance or similarities between pairs of objects. Groups are formed from individual entities by combining the shortest distances or the greatest similarities [2]. At first, we must find the shortest distance in $D = \{dik\}$ and combine the corresponding objects, for example, U and V, to get a cluster (UV) The steps of the algorithm above the distance between (UV) and the other cluster W is determined in Eq. (3).

$$d_{(UV)W} = maks[d_{UW}, d_{VW}] \quad (3)$$

Information :

d_{uv} = the distance between closest neighbors of clusters U and W

d_{vw} = the distance between the closest neighbors of clusters V and W

D. Process Modelling

The process model is a process flow that supports the activities of a company in its operations [13]. Process models can be drawn manually by analysts or can be generated based on event logs [3]. An event log is a record from the process that contains activity case on sequence data and gets from executed in an information system. In addition, it is an important process because of its strength to predict the time, cost, and used the resource. Knowing the benefits of event logs it is necessary to understand the behavior in the event logs so that they can be model to get a better looking for analysis [10].

E. Performance Evaluation Measure

Performance Evaluation Measure (PEM) is one stage bundle used to measure the performance of a system. PEM in many cases is used in training data, replacement to determine the model that has been made. There are many calculations to get the value of PEM, usually applied as a combination or also partially [14]. Some Calculations in PEM include

1) Precision.

Precision is the level of accuracy between user requests and answers system.

2) Accuration.

Accuration is a comparison between information answered by the system correctly with all information

3) Recall.

A recall is a measure of accuracy between the same information as information had been called before.

Formula is determined in Eq. (3) until (5).

$$\text{P}recision = \frac{TP}{FP + TP} \quad (3)$$

$$Accuration = \frac{TN + TP}{FN + FP + TN + TP} \quad (4)$$

$$\text{Re}\,call = \frac{TP}{FN + TP} \quad (5)$$

Information :
TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

III. METHOD

The grouping analysis process diagram uses k-means cluster method, hierarchical complete linkage and hierarchical single linkage can be seen in Figure 1 below.

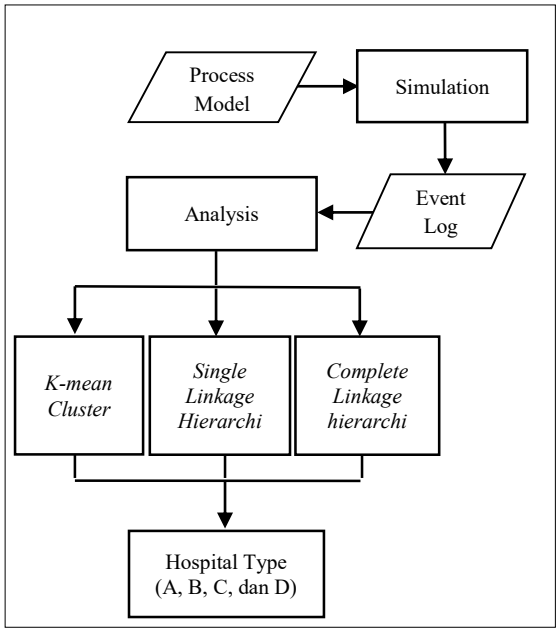

FIGURE I. HOSPITAL PROCESS MODEL DIAGRAM

Based on Figure I above the process model in this study uses a "Process Maker" application where the process is formed in the form of a process model in the hospital both when the patient enters or exits. The process model is used to measure a company's process performance and show the problems associated with the process. The model illustrates the entire process in the hospital system, well-maintained, inpatient, emergency, etc. Then this process model is transformed into information. To make the process model into information, a simulation is made on the process model that has been made. First, this experiment forms a log of processes in the hospital instantaneously using a workflow system. Event logs are used to analyze and find problems in business processes. The log used is based on hospital criteria.

Simulations were carried out in 40 hospitals with each of the 10 patients treated well, inpatient, and emergency department. The results of the simulation are in the form of event log which will later be translated into data. The data is taken based on hospital indicators that have been determined and then analyzed using k-mean cluster analysis, complete hierarchy, and single linkage hierarchy. Based on the existing results, the accuracy, precision, and recall of each analysis will be compared based on which hospital classification provides the best accuracy, precision, and recall.

IV. RESULT AND ANALYSIS

Data from 40 hospitals resulting from process model simulations are based on several predetermined indicators is explained in Table I follows.

Table I. Data Simulation 40 Hospitals

Number	C	I_2	I_3	I_4	I_5	I_6	I_7
1	A	23	7	6	1	71	12
2	A	22	7	7	1	85	18
3	B	12	4	4	1	74	43
4	C	9	2	2	1	99	24
5	B	15	4	4	1	79	38
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
38	C	9	3	3	1	116	56
39	D	8	1	1	0	89	55
40	B	13	4	4	1	111	34

Information:
C = hospital classification
I_1 = number of doctors
I_2 = number of pharmacists in outpatient room
I_3 = number of pharmacists in the inpatient room
I_4 = number of pharmacists in the emergency room
I_5 = waiting time in the outpatient room
I_6 = waiting time for prescription medication

A. Characteristic Data

The following are the characteristics of the simulation data of 40 houses pain is explained in Table II below.

Table II. Simulation Characteristics Data of Human Resources Indicators

Variable	Mean	Varian	Min	Median	Max
X_2	14.9	35.22	5	13.5	25
X_3	4.5	7.23	1	4	10
X_4	4.475	7.38	1	4	10
X_5	0.085	0.13	0	1	1

The results of Table II above can be seen that the average of the simulation data of 40 hospitals is for indicators of human resources with a variable number of 14.9 or 15 people. Data diversity is 35.22 or 36 people with the least number of doctors in the hospital are 4 people and the number of doctors is at most 25 people.

Characteristics of time indicators from data from 40 hospital simulation data are explained in Table III as follows.

Table III. Simulation Characteristic Data of Time Indicators

Variabel	Mean	Varian	Min	Median	Max
X_6	89.08	313.97	60	90.5	119
X_7	30.9	222.5	12	27.5	58

Based on Table III above, it can be estimated that the average simulation data of 40 hospitals for time indicators with outpatient waiting time variables is 89.08 or 90 minutes. The minimum amount of waiting time for outpatients at the hospital is 60 minutes and at most 119 minutes

B. Results of The K-Means Cluster Method, Hierarchical Complete Linkage and Hierarchical Single Linkage

1) K-means Cluster Method

The K-Means Cluster method is a clustering method which groups all data into k clusters, where the value of k is predetermined. This study establishes 4 clusters that will be formed. The following is a grouping of cluster members using the k-mean cluster method is explained in Table IV.

Table IV. K-Means Cluster Grouping Results

Cluster	Hospital Type	List of Hospitals	Total Hospitals
1	A	R 1, R 2, R 6, R 9, R 10, R 13, R 15, R 19, R 21, R 26, R 30	11
2	B	R 8, R 14, R 18, R 20, R 24, R 29, R 32, R 38, R 39, R 40	10
3	C	R 3, R 5, R 7, R 12, R 17, R 22, R 25, R 34	8
4	D	R 4, R 11, R 16, R 23, R 27, R 28, R 31, R 33, R 35, R 36, R 37	11
Total			40

Information :
R 1…40 = Hospitals

The results of the k-mean cluster analysis in Table IV show that cluster 1 is categorized as type A hospital consisting of 11 hospitals namely hospital 1, hospital 2, hospital 6, hospital 9, hospital 10, hospital 13, hospital 15, hospital 15, hospital 19, hospital 21, hospital 26, hospital 30, cluster 2 with type B consisting of 10 hospitals namely hospital 8, hospital 14, hospital 18 , hospital 20, hospital 24, hospital 29, hospital 32, hospital 38, hospital 39, hospital 40, cluster 3 type C consisted of 8 hospitals namely hospital 3, hospital 5, hospital 7, hospital 12, hospital 17, hospital 22, hospital 25, hospital 34, and cluster 4 type D consisted of 11 hospitals namely hospital 4, hospital 11, hospital 16, hospital 23, hospital 27, hospital 28, hospital 31, hospital 33, hospital 35, hospital 36, hospital 37.

2) Hierarchical Complete Linkage

The hierarchical complete linkage is formed with do an approach without first determining the number of groups and determining distances are in the longest distance (smallest similarity) with each other. The following is a grouping of hospital indicators using the hierarchy complete linkage method to determine the classification of hospitals is explained in Table V.

Table V. Hierarchical Complete Linkage Grouping Results

Cluster	Hospital Type	List of Hospitals	Total Hospitals
1	A	R 1, R 10, R 15, R 19, R 30	5
2	B	R 2, R 4, R,6, R 9, R 11, R 16, R 21,R 23, R 26, R 27, R 28,R 31, R 33, R 35, R 36, R 37	16
3	C	R 3, R 5, R 7, R 8, R 12, R 17, R 22, R 25, R 34, R 39	10
4	D	R 14, R 18, R 20, R 24, R 29, R 32, R 38, R 40	8
Total			40

Based on the hierarchical complete linkage analysis in Table V shows that cluster 1 is categorized as type A hospital consisting of 5 hospitals namely hospital 1, hospital 10, hospital 15, hospital 19, hospital 30, cluster 2 with type B consisting of 16 hospitals namely hospital 2, hospital 4, hospital 6, hospital 9, hospital 11, hospital 16, hospital 21, hospital 23, hospital 26, hospital 27, hospital 28, hospital 31, hospital 33, hospital 35, hospital 36, hospital 37, Clusters 3 type C consisted of 10 hospitals namely hospital 3, hospital 5, hospital 7, hospital 8, hospital 12, hospital 17, hospital 22, hospital 25, hospital 34, hospital 39, and cluster 4 type D consisted of 8 hospitals namely hospital 14, hospital 18, hospital 20, hospital 24, hospital 29, hospital 32, hospital 38, hospital 40.

3) Hierarchical Single Linkage

The hierarchical single linkage is formed with conduct discussions without determining the number of groups first and groups forming groups with short groups. The following is a grouping of hospital indicators using the single hierarchy linkage method to determine hospital classification is explained in Table VI.

Table VI. Hierarchical Single Linkage Grouping Results

Cluster	Hospital Type	List of Hospitals	Total Hospitals
1	A	R 1, R 2, R 3, R 5, R 7, R 9, R 10, R 12, R 13, R 15, R 17, R 19, R 21, R 22, R 25, R 30, R 34, R 36,	18
2	B	R 4, R 6, R 11, R 16, R 23, R 26, R 27, R 28, R 31, R 33, R 35, R 37	12
3	C	R 8, R 14, R 29, R 32, R 39	5
4	D	R 18, R 20, R 24, R 38, R 40	5
Total			40

Based on hierarchical single linkage analysis in Table VI. showed that cluster 1 was categorized as type A hospital consisting of 18 hospitals namely hospital 1, hospital 2, hospital 3, hospital 5, hospital 7, hospital 9, hospital 10, 12 hospitals, hospital 13, hospital 15, hospital 17, hospitals 19, hospital 21, hospital 22 , hospital 25, hospital 30, hospital 34, hospital 36, cluster 2 with type B consisted of hospital 12, hospital 4 , hospital 6, hospital 11, hospital 16, hospital 23, hospital 26, hospital 27, hospital 28, hospital 31, hospital 33, hospital 35, hospital 37, cluster 3 type C consisted of 5 hospitals namely hospital 8, hospital 14, hospital 29, hospital 32, hospital 39, and cluster 4 type D classes consisted of 5 hospitals namely hospital 18, hospital 20, hospital 24, hospital 38, hospital 40.

C. Confusion Matrix Table

The following is a data prediction and reality using the k-means cluster, hierarchical complete linkage and hierarchical single linkage method can be seen in below.

1) K-means Cluster

The following predictions and reality using k-means cluster can be explained in table VII below.

Table VII. Prediction And Reality using K-Means Cluster

		Reality				
		A	B	C	D	Total
Prediction	A	7	1	2	3	13
	B	1	5	5	3	14
	C	1	1	1	4	7
	D	2	3	0	1	6
Total		11	10	8	11	40

Based on Table VII above, we find that accuracy, accuracy, memory are as follows in Table VIII below.

Table VIII. Acuration, Prediction and Recall using K-Means Cluster

	A	B	C	D	Total
True Positive (TP)	7	5	1	1	14
False Positive (FP)	4	5	7	10	26
False Negative (FN)	6	9	6	5	26
True Negative (TN)	23	21	26	24	94

$$Acuration = \frac{TP + TN}{TP + FP + FN + TN} = 67.5\%$$

Based on accuracy for the k-means cluster method has a classification accuracy of 67.5% and the recall can be explained in Table IX below.

Table IX. Result Recall using K-Means Cluster

Recall	Formula	Average
A	$TP(A)/Total$	0.175
B	$TP(B)/Total$	0.125
C	$TP(C)/Total$	0.025
D	$TP(D)/Total$	0.025
Average		0.0875

So based on the k-mean cluster method the value of recall is 8.75% and the prediction can be explained in Table. X below.

Table X. Result Prediction using K-Means Cluster

Precision	**Formula**	**Average**
A	$TP(A)/Tot.predic(A)$	0.54
B	$TP(B)/Tot.predic(B)$	0.36
C	$TP(C)/Tot.predic(C)$	0.14
D	$TP(D)/Tot.predic(D)$	0.17
	Average	0.30

Based on the k-mean cluster method the value of proximity (precision) is 30%.

2) Hierarchical Complete Linkage

The following predictions and reality using hierarchical complete linkage can be explained in Table XI below.

Table XI. Prediction And Reality using Hierarchical Complete Linkage

		Reality				
		A	B	C	D	Total
Prediction	A	4	1	0	1	6
	B	6	3	5	2	16
	C	2	5	1	2	10
	D	1	5	1	1	8
Total		13	14	7	6	40

Based on the above table, then find the accuracy, precision, recall as follows in the Table XII below.

Table XII. Acuration, Prediction and Recall using Hierarchical Complete Linkage

	A	B	C	D	Total
True Positive (TP)	4	3	1	1	9
False Positive (FP)	9	11	6	5	31
False Negative (FN)	2	13	9	7	31
True Negative (TN)	25	13	24	27	89

$$Acuration = \frac{TP + TN}{TP + FP + FN + TN} = 61.25\%$$

So Accuracy for hierarchical complete linkage method has classification accuracy of 61.25% and recall be explained in the Table XIII below

Table XIII. Result Recall using Hierarchical Complete Linkage

Recall	Formula	Average
A	$TP(A)/Total$	0.25
B	$TP(B)/Total$	0.075
C	$TP(C)/Total$	0
D	$TP(D)/Total$	0.025
Average		0.0875

So based on the hierarchical complete linkage method the value of is 8 recall 8.75% and prediction can be explained in Table XIV below.

Table XIV. Result Prediction using Hierarchical Complete Linkage

Presisi	Formula	Average
A	$TP(A)/Tot.predic(A)$	0.56
B	$TP(B)/Tot.predic(B)$	0.25
C	$TP(C)/Tot.predic(C)$	0
D	$TP(D)/Tot.predic(D)$	0.2
	Average	0.25

Based on the hierarchical complete linkage method the value of proximity (precision) is 25%

1) Hierarchical Single Linkage

The following predictions and reality using hierarchical single linkage can be explained in Table XV below.

Table XV. Prediction and Reality using Hierarchical Single Linkage

		Reality				
		A	B	C	D	Total
Prediction	A	10	6	1	1	18
	B	2	3	5	2	12
	C	1	2	0	2	5
	D	0	3	1	1	5
Total		13	14	7	6	40

Based on the table above, then we find the accuracy, precision, recall as follows in the Table XVI below.

Table XVI. Result Recall using Hierarchical Single Linkage

	A	B	C	D	Total
True Positive (TP)	10	3	0	1	14
False Positive (FP)	3	11	7	5	26
False Negative (FN)	8	9	5	4	26
True Negative (TN)	19	17	28	30	94

$$Acuration = \frac{TP + TN}{TP + FP + FN + TN} = 67.5\%$$

So accuracy for the hierarchical single linkage method has classification accuracy of 67.5% and the recall can be explained in the Table XVII below.

Table XVII. Result Recall using Hierarchical Single Linkage

Recall	Formula	
A	$TP(A)/Total$	0.1
B	$TP(B)/Total$	0.075
C	$TP(C)/Total$	0.025
D	$TP(D)/Total$	0.025
	Average	0.056

So based on the hierarchical single linkage method the value of proximity (precision) is 5.6% and the prediction can be explained Table XVIII below.

Table XVIII. Result Prediction using Hierarchical Single Linkage

Presisi	Formula	Average
A	$TP(A)/Tot.predic(A)$	0.67
B	$TP(B)/Tot.predic(B)$	0.19
C	$TP(C)/Tot.predic(C)$	0.1
D	$TP(D)/Tot.predic(D)$	0.13
	Average	0.27

So based on hierarchical single linkage method, the value of proximity (precision) is 27%.

V. CONCLUSION

Hospitals are institutions that provide health services both in outpatient rooms, inpatients and emergency care. Hospitals in Indonesia are classified into 4 classes, namely classes A, B, C, and D. The classification is based on predetermined indicators such as the number of doctors, the number of pharmacists and several indicators. Determination of these indicators, then analyzed using k-means clusters, hierarchical complete linkages, and hierarchical single linkages to get classes from the hospital. The results of the analysis can be seen that the calculation of accuracy, precision, and recall that have been done using the k-means cluster method when compared with 2 methods, namely hierarchical complete linkages, and hierarchical single linkages have better values of accuracy, precision, and recall to get the class from the hospital.

REFERENCES

[1] F. E. B. Setyawan and S. Supriyanto,"Hospital Management.". Sidoarjo, Zifatama Jawara, 2019.

[2] H. Dermawan, R. Sarno, A. S. Ahmadiyah, K.R. Sungkono, C.S. Wahyuni, "Anomaly Detection Based on Control-flow Pattern of Parallel Business Processes", Telecommunication, Computing, Electronics and Control (TELKOMNIKA), Vol. 16 No. 6, pp.2808-2815, 2018. DOI: 10.12928/TELKOMNIKA

[3] L. Hakim, R. Sarno, K. R. Sungkono," Modified Alpha++ Algorithm for Discovering the Hybrid of Non-Free Choice and Invisible Task of Business Processes", in International Journal of Intelligent Engineering and Systems, Vol. 12 No. 3, 2019. DOI: 10.22266/ijies2019.0630.04.

[4] R. Sarno, K.R. Sungkono, R. Johanes, D. Sunaryono, "Graph-Based Algorithms for Discovering a Process Model Containing Invisible Tasks", in International Journal of Intelligent Engineering and Systems, Vol. 12 No. 2, 2019. DOI: 10.22266/ijies2019.0230.09.

[5] A. Burratin,"Learning Process Modeling Phases From Modeling Interactions And Eye Tracking Data", ELSEVIER, 2018

[6] Y. A. Effendi and R. Sarno,"Implementation of the Semantic Web in Business Process Modeling Using Petri Nets", Lab-Based Education Enterprise Intelligent System, 2018.

[7] P. Franti and S. Sieranoja," How Much Can K-means be Improved by using Better Initialization and Repeat", ELSEVIER, 95-112, 2019.

[8] C. Siregar,"Hospital Pharmacy(Theory and Application)" , The EGC Medical Book, pp 13-15, 2004

[9] D. A. Rahmawati, L. N. Aini, R. Sarno, C. Fatichah, D. Sunaryono," Lab-Based Education Enterprise Intelligent System, 2017.

[10] Maryamah, R. Sarno, A. L. Nurlaili," Determining Linear Temporal Logic Formula for Decomposed Process Model,"Comparison of Behavioral Similarity use TARs and Naïve Algorithm for Calculating Similarity in Business Process Model", in 3rd International Conference on Science in Information Technology(ICSITech), 2017. DOI: 10.1109/icsitech.2017.8257095

[11] A. Imron,"Analisis Sentimen Terhadap Tempat Wisata Di Kabupaten Rembang Menggunakan Metode Naive Bayes Classifier". Tugas Akhir Universitas Islam Indonesia, 2019.

[12] V. C. Addad, V. Kanade, F. M. Trenn, C. Mathieu,"Hierarchical Clustering: Objective Functions and Algorithms",SIAM, 2017. arXiv:1704.02147v1

[13] D. Suhendra. Implementasi dan Analisis Akurasi Pengenalan Objek Gambar Menggunakan Metode Complete-Link Clustering. *IEEE Paper Templet in A4 Vol 1*, 180, 2018.

[14] R. Subekti, R. Kusumawati, E. R. Sari," K-means Clustering and Average Linkage dalam Pembentukan Portofolio Saham", Seminar Matematika dan Pendidikan Matematika UNY, PT-220, 2017.