

IMPLEMENTATION OF CLUSTERING WITH THE K-MEANS ALGORITHM METHOD IN HOSPITAL GROUPING IN WEST JAVA

Nurhaeningsih*¹, Denni Pratama²

^{1,2} Sekolah Tinggi Manajemen Informatika dan Komputer IKMI Cirebon, Indonesia

Accepted: 03rd April 2023. Approved: 30st April 2023. Published: 30th April 2023.

ABSTRACT

Hospitals as health service institutions must provide services to the community in the local and international scope. Based on this, in the last few decades the term accreditation has emerged to assess the quality of an organization including hospitals. The state has an obligation to meet the public's need for health services which is regulated in the 1945 Constitution of the Republic of Indonesia, as well as supporting laws and regulations, namely Law Number 25 concerning Public Services, Law Number 36 of 2009 concerning health, Law Number 44 of 2009 concerning Hospitals, and Law 37 of 2008 concerning the Ombudsman of the Republic of Indonesia. to find out about matters relating to hospital accreditation and patient safety in order to improve the quality of health services in hospitals. The research design used is data mining. With clustering algorithm. Data taken from [opendata.jabarprov](https://opendata.jabarprov.go.id/id/dataset/register-rumah-saki-berdasarkan-tipe-dan-class-di-jawa-barat) <https://opendata.jabarprov.go.id/id/dataset/register-rumah-saki-berdasarkan-tipe-dan-class-di-jawa-barat>. In data mining, the results show that hospital accreditation in Indonesia is carried out to assess hospital compliance with accreditation standards. Hospital accreditation is linked to quality assessment. However, actually the quality itself is the outcome of the accreditation implementer, while accreditation only assesses whether the service has met the standard or not without measuring the quality of the service. Hospital accreditation is an acknowledgment that has been recognized by the government for hospitals because they have met predetermined standards.

Research Paper

INOVASIA

Keywords: Data mining, Clustering, Hospital accreditation.

INTRODUCTION

The definition of a hospital based on the Decree of the Minister of Health of the Republic of Indonesia No.983/Menkes/SK/XI/1992 states that a public hospital is a hospital that has provided basic, specialist and sub-specialistic health services. The hospital is a health service institution that is complex, expert-dense, and capital-intensive. This complexity arises because the hospital's services involve various service, education and research functions, and cover various levels and types of discipline, so that the hospital is able to carry out professional functions both in the medical technical and health administration fields. to maintain and improve the quality of the hospital must have a measure that guarantees quality improvement at all levels (Nisak, 2020).

Law no. 44 of 2009 concerning Hospitals requires hospitals to undergo accreditation. Thus hospitals must implement hospital accreditation standards, including other standards that apply to hospitals in accordance with the elaboration in the 2011 edition of the Hospital Accreditation Standards (RI, 2022).

The hospital is one of the government infrastructures that functions to carry out basic health efforts or referral health and supporting health efforts that are complex, this is because hospital services consist of various service functions, education, research, and cover various levels and types of scientific disciplines (Suhartina, 2019).

Hospital statistics are a collection of data in the form of numbers that describe information on the condition of a hospital that

*Correspondance Address

E-mail: nurhaeningsih@gmail.com

has been used for decision making and is sourced from medical record data (Nisak, 2020).

Hospital accreditation is a process in which an institution, which is independent, conducts an assessment of the hospital. The goal is to determine whether the hospital meets the standards that will be designed to improve the safety and quality of care in the hospital. The accreditation standard is in the form of an optimal and easy to achieve requirement. Accreditation shows a hospital's real commitment to improving the safety and quality of patient care, ensuring that the service environment is safe and the hospital always strives to reduce risks for patients and hospital staff. Thus accreditation is needed as an effective way to evaluate the quality of a hospital, which also acts as a management tool (RI, 2022).

LITERATURE REVIEW

Hospital grouping is an important aspect of health service management. In this grouping, hospitals will be grouped based on their characteristics, performance, and service capabilities (Nisak, 2020). The clustering method is an option in hospital grouping, where this method can group hospitals based on similar characteristics.

One of the most frequently used clustering methods is the k-means method. The k-means method is used in various fields, such as customer clustering, gene clustering in bioinformatics, and medical data clustering. In the context of hospital clustering, several studies have been conducted previously to apply the k-means method.

Song et al. (2013) used hospital performance data in South Korea to classify hospitals into several groups based on the performance capabilities of each hospital. The results showed that the k-means method can be used to classify hospitals effectively. Another study conducted by Sano et al. (2016) used the hierarchical clustering method to classify hospitals in Turkey based on patient profiles and hospital characteristics.

In this study, the authors implemented the k-means method to classify hospitals in the West Java region based on the characteristics and performance of each hospital. This

method is expected to assist in making strategic decisions to improve the quality of health services in the West Java region.

In this study, the authors used data consisting of 52 hospitals in the West Java region. The data used consists of 12 variables which include the number of doctors, the number of nurses, the number of beds, the ratio of doctors per bed, the ratio of nurses per bed, and so on. After data processing and analysis using the k-means method, hospitals in the West Java region were successfully grouped into 4 groups.

The results of this study indicate that the clustering of hospitals using the k-means method can assist in making strategic decisions and planning the development of health services. In addition, the use of the k-means method can increase the efficiency and effectiveness of hospital management and improve the quality of health services in the West Java region.

METHODS

This research data takes secondary data, namely data sources that do not directly provide data to data collectors, for example through other people or through documents. In this study, the source of secondary data is in accordance with the Labor Law, books, journals, articles related to the research topic regarding internal control systems over payroll systems and procedures in an effort to support labor cost efficiency (STEI Indonesia, 2017).

Examples of secondary data are data obtained from most of the results of research that was previously carried out. For example, such as data obtained from sources such as books, websites, or government documents. This data is taken from opendata.jabarpovr. In collecting data related to research, the authors use secondary data collection techniques to obtain specific results.

RESULTS AND DISCUSSION

Data Collection Techniques

This research data takes secondary data, namely data sources that do not directly provide data to data collectors, for example through other people or through documents. In this research, the source of secondary data

is in accordance with the Labor Law, books, journals, articles related to the research topic regarding internal control systems over payroll systems and procedures in an effort to support labor cost efficiency.

Design Stages

The design stage used is Knowledge Discovery in Databases (KDD) which is a whole non-trivial process for finding and

identifying patterns in the data, where the patterns found are valid, new, useful and understandable. KDD relates to integration techniques and scientific discoveries, interpretation and visualization of patterns in a number of data sets (Khatisma, 2014).

The following is an image of the final project design that will be applied:

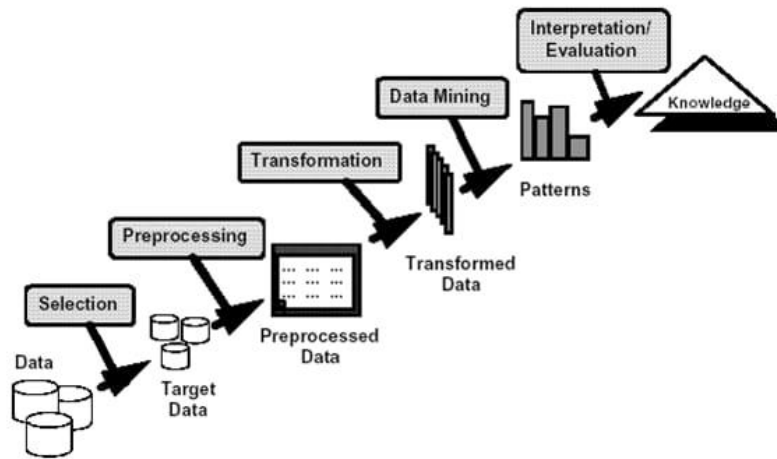


Figure 1 KDD process stages

In the design or research stage, the authors use the knowledge discovery in database (KDD) stage, the stages are as follows:

Data Selection

Data selection is the process of retrieving data related to analysis from the

database. At this stage, the technique of obtaining a reduced representation of the data and minimizing the loss of data information is carried out. This includes attribute reduction and data compression methods.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|----|---------------|---------------|---------------------|---------------------|----------------------|---------------------|-------------------|-------|---|---|---|
| 1 | id | kode_provinsi | nama_provinsi | kode_kabupaten_kota | nama_kabupaten_kota | nama_rumah_sakit | jenis_rumah_sakit | kelas_rumah_sakit | tahun | | | |
| 2 | 1 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS PARIU DR. M. GOEN | RUMAH SAKIT KHUSUS | B | 2019 | | | |
| 3 | 2 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DAERAH CII | RUMAH SAKIT UMUM | B | 2019 | | | |
| 4 | 3 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DAERAH CI | RUMAH SAKIT UMUM | B | 2019 | | | |
| 5 | 4 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM BINA HUSA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 6 | 5 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS IBU DAN ANAK CIT | RUMAH SAKIT IBU DAN | C | 2019 | | | |
| 7 | 6 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM MARY CILEI | RUMAH SAKIT UMUM | C | 2019 | | | |
| 8 | 7 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM MH. THAMI | RUMAH SAKIT UMUM | C | 2019 | | | |
| 9 | 8 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DR MUHAV | RUMAH SAKIT UMUM | C | 2019 | | | |
| 10 | 9 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM TRIMITRA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 11 | 10 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM FAMILY ME | RUMAH SAKIT UMUM | C | 2019 | | | |
| 12 | 11 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DAERAH LE | RUMAH SAKIT UMUM | B | 2019 | | | |
| 13 | 12 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM CITAMA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 14 | 13 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM ANNISA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 15 | 14 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM SENTRA ME | RUMAH SAKIT UMUM | B | 2019 | | | |
| 16 | 15 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS SENTOSA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 17 | 16 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DAERAH CII | RUMAH SAKIT UMUM | C | 2019 | | | |
| 18 | 17 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM DR. SISMAE | RUMAH SAKIT UMUM | C | 2019 | | | |
| 19 | 18 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM HERMINA N | RUMAH SAKIT UMUM | C | 2019 | | | |
| 20 | 19 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS IBU DAN ANAK ASI | RUMAH SAKIT IBU DAN | C | 2019 | | | |
| 21 | 20 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS IBU DAN ANAK KEI | RUMAH SAKIT IBU DAN | C | 2019 | | | |
| 22 | 21 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM PERMATA J | RUMAH SAKIT UMUM | C | 2019 | | | |
| 23 | 22 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM KARYA BHA | RUMAH SAKIT UMUM | C | 2019 | | | |
| 24 | 23 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS UMUM RUMAH SEI | RUMAH SAKIT UMUM | C | 2019 | | | |
| 25 | 24 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | RS EMC SENTUL | RUMAH SAKIT UMUM | B | 2019 | | | |

Figure 2 Data Selection

Of the 9 fields or attributes, namely id, province code, province name, district code, district/city name, hospital name, hospital type, hospital class, year.

1. Pre-processing/Cleaning
 - a. Pre-processing and data cleaning are basic operations such as noise removal.
 - b. Before the data mining process can be carried out, it is necessary to carry out a

cleaning process on the data that is the focus of KDD.

- c. The cleaning process includes removing duplicate data, checking inconsistent data, and correcting errors in data.
- d. Enrichment process is carried out, namely the process of enriching existing data with other (external) data or information.

| Row No. | kode_provinsi | nama_provinsi... | kode_kabup... | nama_kabu... | nama_ruma... | jenis_rumah... | kelas_ruma... | tahun |
|---------|---------------|------------------|---------------|---------------|----------------|----------------|---------------|-------|
| 1 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS PARU DR... | RUMAH SAKI... | B | 2019 |
| 2 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM DA... | RUMAH SAKI... | B | 2019 |
| 3 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM DA... | RUMAH SAKI... | B | 2019 |
| 4 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM BL... | RUMAH SAKI... | C | 2019 |
| 5 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS IBU DAN ... | RUMAH SAKI... | C | 2019 |
| 6 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM MA... | RUMAH SAKI... | C | 2019 |
| 7 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM M... | RUMAH SAKI... | C | 2019 |
| 8 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM D... | RUMAH SAKI... | C | 2019 |
| 9 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM TR... | RUMAH SAKI... | C | 2019 |
| 10 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM FA... | RUMAH SAKI... | C | 2019 |
| 11 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM DA... | RUMAH SAKI... | B | 2019 |
| 12 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM CL... | RUMAH SAKI... | C | 2019 |
| 13 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM AN... | RUMAH SAKI... | C | 2019 |
| 14 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS UMUM SE... | RUMAH SAKI... | B | 2019 |
| 15 | 32 | JAWA BARAT | 3201 | KABUPATEN ... | RS SENTOSA | RUMAH SAKI... | C | 2019 |

ExampleSet (738 examples, 0 special attributes, 8 regular attributes)

Figure 3 Pre-processing/Cleaning data

Here the data that is cleaned is the ID of the West Java hospital data.

1. Data Transformation

The integration process on the data that has been selected, so that the data is suitable for the data mining process. It is a process that depends heavily on the type or pattern of information sought in the database, changes to the data type which are changed nominally to nominal (0-1) Coding is a process of transforming the selected data so that the data is suitable for the data mining process.

The coding process in KDD is a creative process and really depends on the type or pattern of information to be searched for in the database.

2. Data mining

Data mining is the analysis of data to find clear relationships and conclude that were not previously known in an easy-to-understand and useful way for the owner of the data (Yuli Mardi, 2019).

Data mining is divided into several groups based on the tasks that can be carried out, namely Description, Estimation,

Prediction, Classification, Clustering, and Association (Hermanto & Muhyidin, 2020).

Data mining is the process of analyzing different data and concluding them into new and important information or knowledge or patterns to increase profits, reduce costs, or even both (Abdullah & Saputra, 2016).

Data mining is a technique that can help predict the phenomenon of an object under study by first knowing or not knowing a particular class of object through input from large data. Data mining is the process of looking for patterns or interesting information in selected data using certain techniques or methods. The k-means method or algorithm in data mining varies greatly, such as characterization, classification, regression, clustering, association, and so on. The selection of the right technique, method or algorithm depends very much on the goals and the overall KDD process.

3. Interpretation/Evaluation

Interpretation/Evaluation is the translation of several patterns resulting from data mining. The resulting pattern of information needs to be presented in a form that is understandable to interested parties.

This stage is part of the KDD process called interpretation. At this stage, it is necessary to check whether the pattern or information found contradicts the facts or hypotheses that existed before. At this stage, the processed data will be evaluated using a performance model to calculate the Davies Bouldin Index (DBI), with the aim of obtaining the most optimal K value. So that it can be concluded whether the hospital data grouping has reached the desired goal.

4. K-Means results on hospital data in West Java

a. Import Data

Open the RapidMiner application that is already installed on the PC. After it opens,

pay attention to the box that says Repository. Then click Import Data after a new window appears, choose the place where the data will be imported. In our example we'll do it in the My Computer location.

b. Data Import Process

After selecting the file in My Computer and selecting import data into Rapidminer, the display will be as above and check the attributes again. The attributes will be displayed as shown in the image. This data has 9 fields or attributes, namely id, province code, name province, district code, district/city name, hospital name, hospital type, hospital class, year.

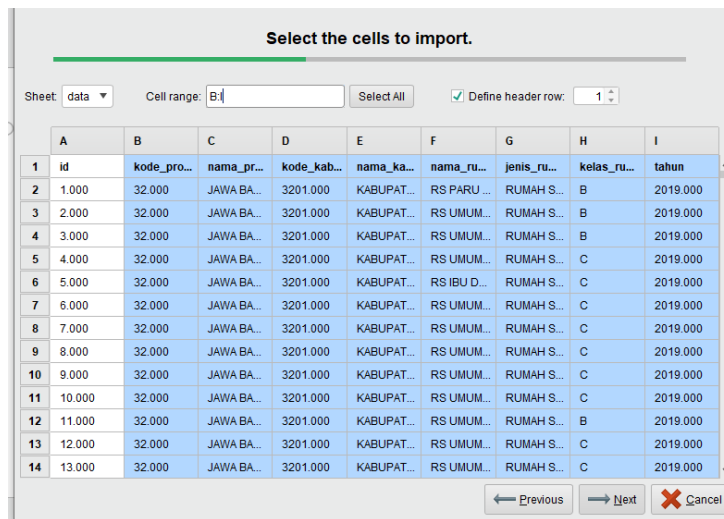


Figure 4 Import data

c. Carry out the Data Cleaning Process

After the data import process, the next step is to carry out the cleaning process on the data, before the data is cleaned, the attributes that will be displayed are provincial code =

integer, province name = polynomial, city district code = integer, city district name = polynomial, hospital name = Polynomial, Hospital Type = Polynomial, Hospital Class = Polynomial, Year = Integer.

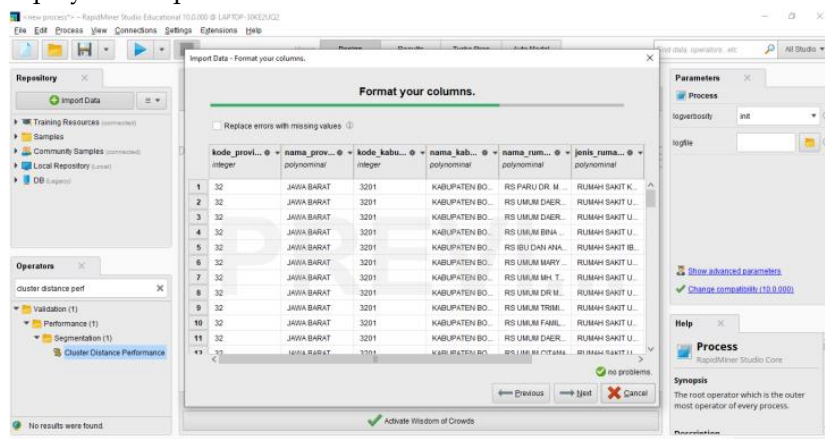


Figure 5 Data Cleaning Process

d. Process Data

After the data has been successfully imported, the next step is to retrieve the data

and enter it into the process sheet page, and click Local Repository. The imported data here is named West Java hospital data then click on data to process in rapidminer.

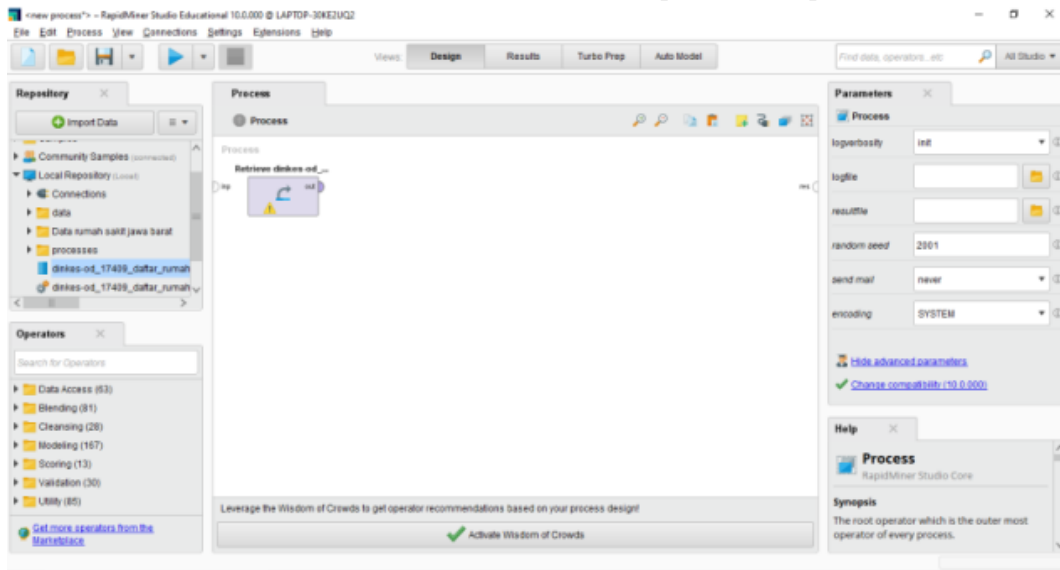


Figure 6 Process Data

e. Enter the Nominal to Numerical Operator in the process

The next step is to enter an Operator named Nominal to Numerical into the process sheet, and the function of the Operator,

namely Nominal to Numerical to change the type from a non-numeric attribute to a numeric attribute type, in Parameters next to the process sheet of the filter type attribute used, namely all and the Coding type used is Unique integers.

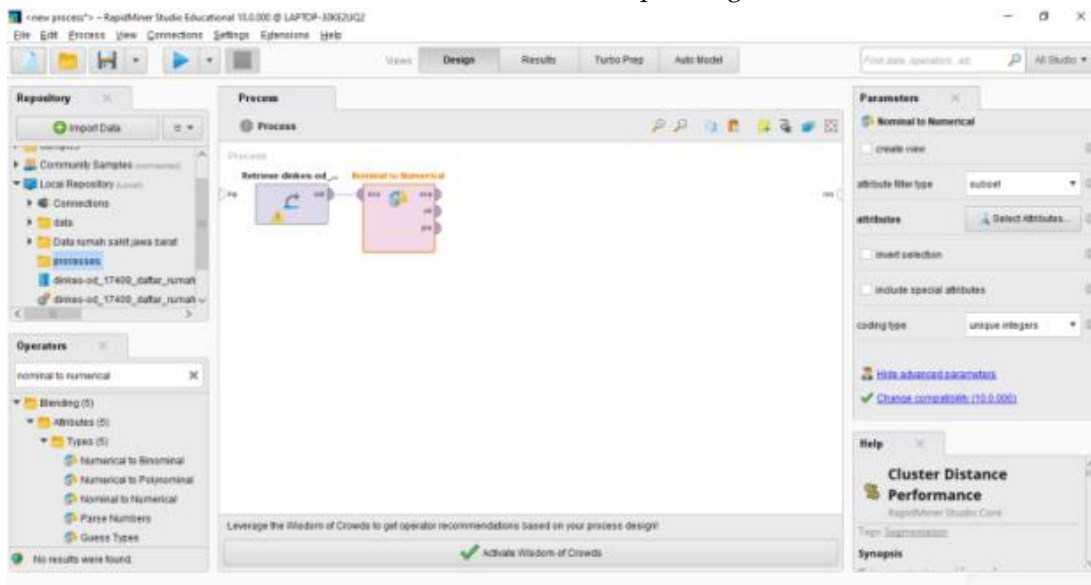


Figure 7 Nominal to Numerical

f. Assign Normalize Operator to the process

After entering the Nominal to Numerical Operator the next step is to enter the Normalize Operator into the process sheet. The function of the Normalize Operator is to

normalize the attribute values so that they are not too high and not too low, in the parameters next to the process sheet. The filter type attribute used is all, The method used is the Range transformation and with a min value of 0.0 and a max value of 0.1.

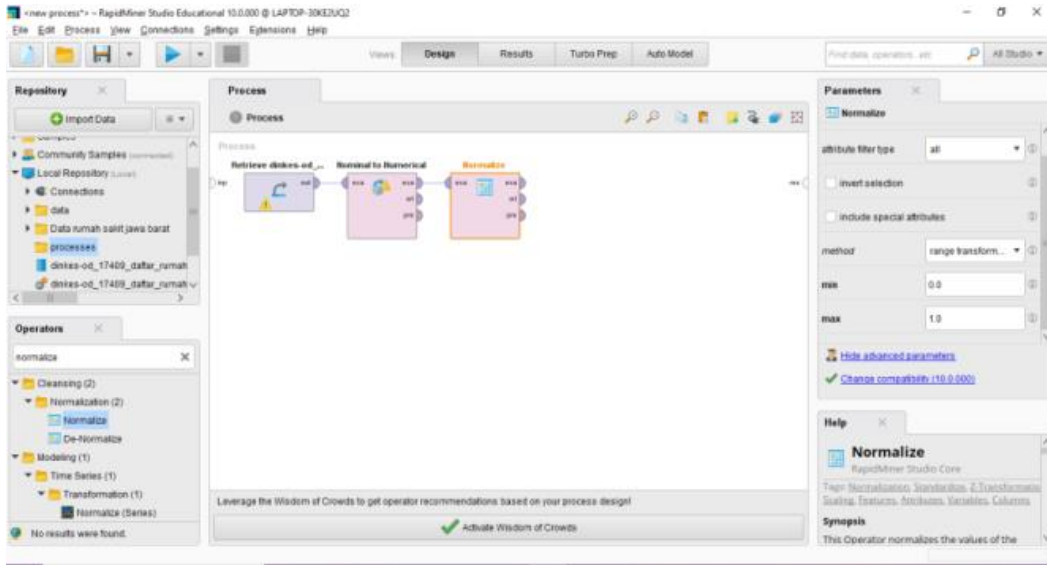


Figure 8 Operator Normalize

- g. Incorporate the Clustering Operator to the process

The next step is how to enter the K-Means Algorithm which is available on the Operator menu in Rapidminer. There will be 4 K-Means options, namely (K-Means, K-Means (H20), K-Means (Karnel), and K-Means (Fast) then choose K-Means after that it will appear

on the process sheet namely Clustering, Then in Parameters there is K meaning Clustering will be grouped into the desired clusters and the run max process means how many times will experiments be carried out, value k = 5 , max runs 10 rounds with measure types of Numerical Divergences and uses the EuclideanDistance method with a max optimization steps of 100.

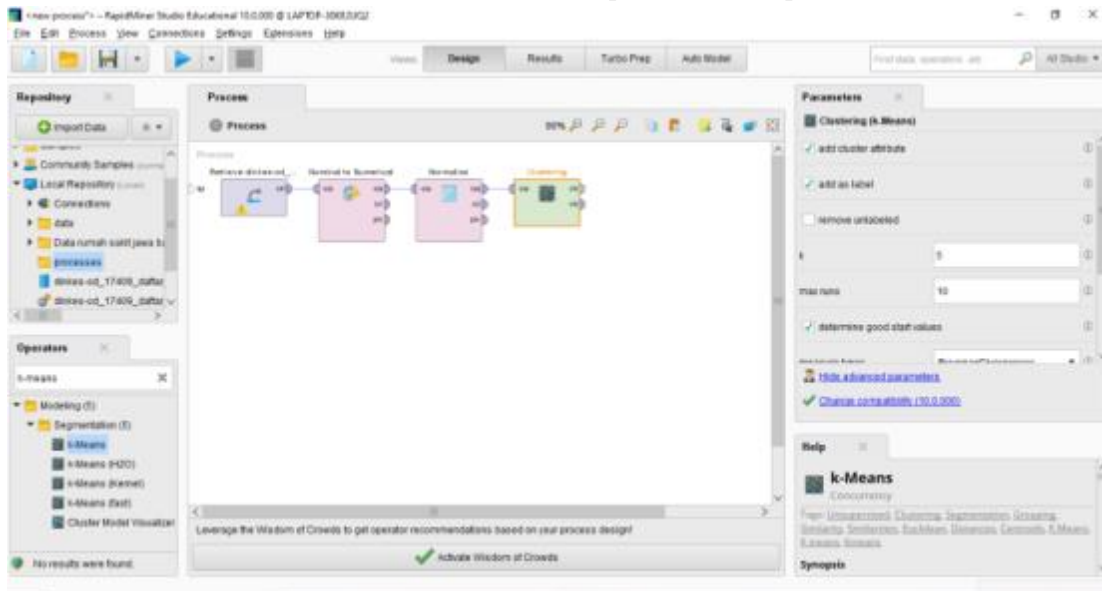


Figure 9 Clustering Operators

- h. Enter Performance (Cluster Distant Performance) to the process

The next step is After entering the Clustering Operator on the process sheet, then entering the Performance Operator (Cluster Distant

Performance) on the process sheet. This operator function is to find out the performance results that will be obtained from the data processing process. In Parameter main criterion select Davies Bouldin and Click Check on Maximize.

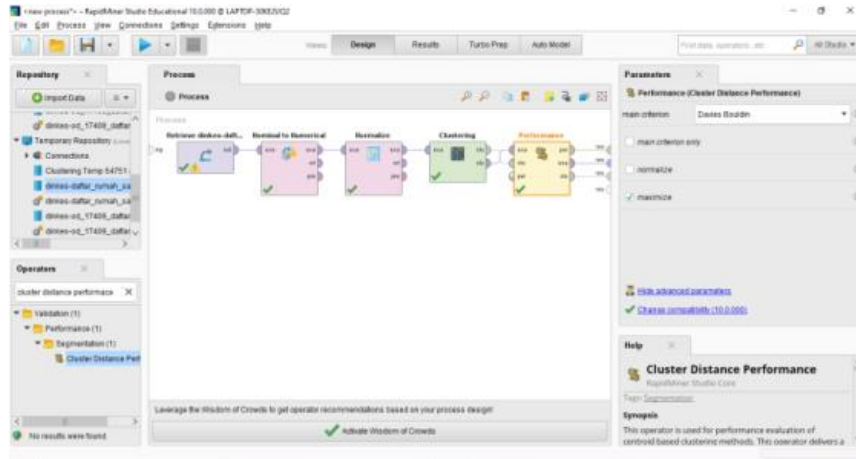


Figure 10 Performance process

i. Stage of the testing process
 After all operators have been entered into the process sheet and make sure the

parameters used are correct and then connect to all operators then click run if all the data is correct then the display on Result will appear as shown below:

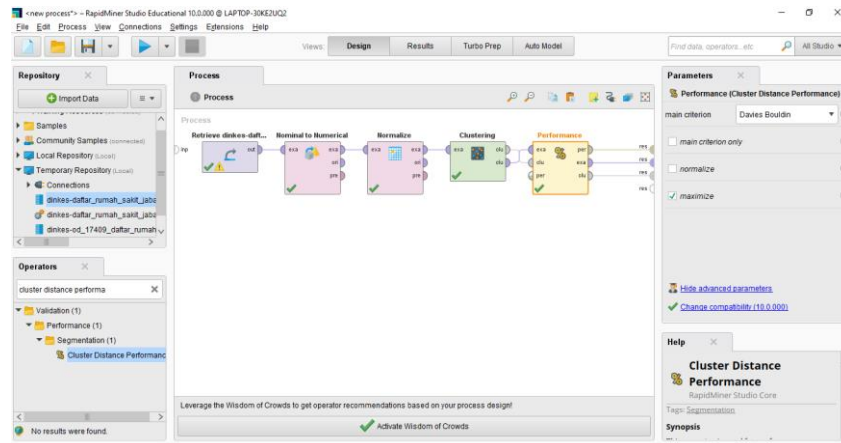


Figure 11 Testing Process

j. Results in Clustering
 The method used is the Clustering Method by applying the K-Means Algorithm where it is known that the value of $k = 5$, max

runs as much as 10 rounds and with the type of measure types is NumericalMeasure using the calculation method namely Euclidean Distance and max optimization step of 100.

| Row No. | id | label | nama_prov... | nama_kab... | nama_ruma... | jenis_rumah... | luas_ruma... | kode_prov... | kodi | |
|---------|----|-----------|--------------|-------------|--------------|----------------|--------------|--------------|------|--|
| 1 | 1 | cluster_2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | |
| 2 | 2 | cluster_2 | 0 | 0 | 0.002 | 0.125 | 0 | 1 | 0 | |
| 3 | 3 | cluster_2 | 0 | 0 | 0.004 | 0.125 | 0 | 1 | 0 | |
| 4 | 4 | cluster_2 | 0 | 0 | 0.006 | 0.125 | 0.250 | 1 | 0 | |
| 5 | 5 | cluster_2 | 0 | 0 | 0.008 | 0.250 | 0.250 | 1 | 0 | |
| 6 | 6 | cluster_2 | 0 | 0 | 0.010 | 0.125 | 0.250 | 1 | 0 | |
| 7 | 7 | cluster_2 | 0 | 0 | 0.012 | 0.125 | 0.250 | 1 | 0 | |
| 8 | 8 | cluster_2 | 0 | 0 | 0.014 | 0.125 | 0.250 | 1 | 0 | |
| 9 | 9 | cluster_2 | 0 | 0 | 0.016 | 0.125 | 0.250 | 1 | 0 | |
| 10 | 10 | cluster_2 | 0 | 0 | 0.018 | 0.125 | 0.250 | 1 | 0 | |
| 11 | 11 | cluster_2 | 0 | 0 | 0.020 | 0.125 | 0 | 1 | 0 | |
| 12 | 12 | cluster_2 | 0 | 0 | 0.022 | 0.125 | 0.250 | 1 | 0 | |
| 13 | 13 | cluster_2 | 0 | 0 | 0.024 | 0.125 | 0.250 | 1 | 0 | |
| 14 | 14 | cluster_2 | 0 | 0 | 0.026 | 0.125 | 0 | 1 | 0 | |

Figure 12 Clustering Results

k. Result of Performance Vector

After the testing process is carried out and produces clusters, the Performance Vector

results will appear with the following images and descriptions of the Performance Vector results:

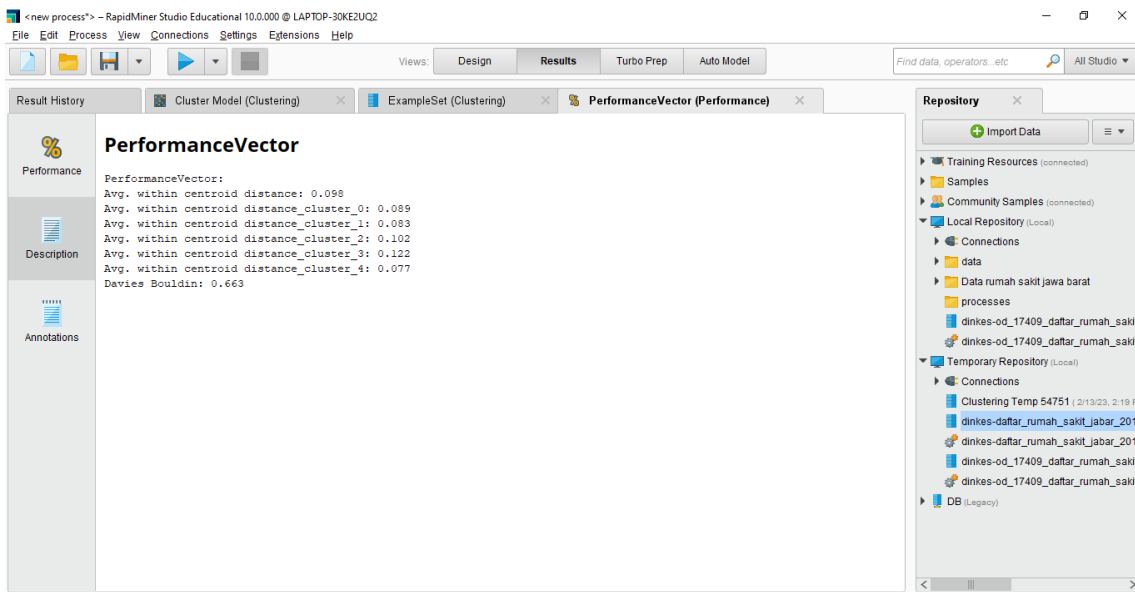


Figure 13 Performance Vector results

l. Davies Bouldin Index results

After carrying out the testing process and producing a Performance Vector, the

results of the Davies Bouldin Index appear with a value of 0.663, which means that it can be called k = optimum.

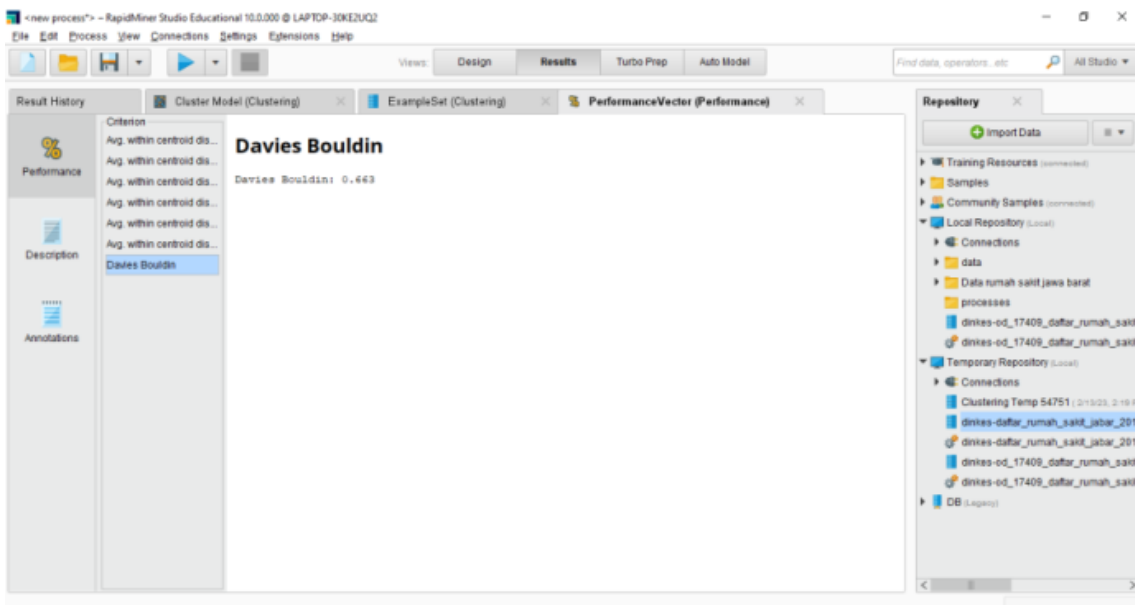


Figure 14 Results of the Davies Boulding Index

m. Cluster Model Results Process

After carrying out the testing process and getting the Performance Vector results, the results of the Cluster Model also appear

that the cluster models produced are Cluster 0 with 165 Items, Cluster 1 with 133 Items, Cluster 2 with 197 Items, cluster 3 with 165 items, cluster 4 with 78 items. The total of all datasets tested is 738 items.

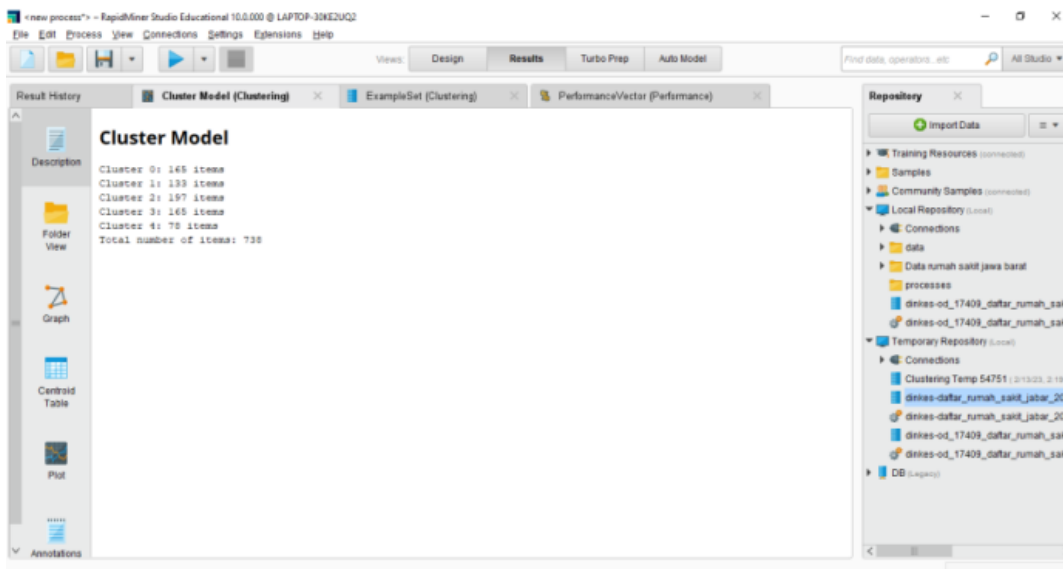


Figure 15 Cluster Models

Cluster Evaluation Results on hospital data in West Java

At this stage the pattern of information that has been generated in the data mining process needs to be displayed in a form that is easy for everyone to understand and understand. Cluster Member Number of Hospital Types and Classes in West Java.

Table 1 Cluster models

| Cluster | Total | Levels |
|--------------|------------------|---------------|
| 0 | 165 items | Moderate (2) |
| 1 | 133 items | little (2) |
| 2 | 197 items | Lots |
| 3 | 165 items | Moderate (1) |
| 4 | 78 items | little (1) |
| Total | 738 items | Levels |

After 5 attempts to determine the result of the k value, in the 5th experiment using the determined k value, the k 5 value produces the optimal Performance Vector (Davies Bouldin Index) value compared to the first, second, third, fourth and so on up to 10 times the run process, the resulting value is a value close to 0 with a value of 0.663.

After carrying out Clustering modeling on hospital data in West Java, the next step is to validate clustering using the cluster distance performance operator so that the DBI (Davies Bouldin Index) value maximizes the distance measurement from one cluster to another cluster. In this operator the parameter

the main criterion is changed to the Davies Bouldin Index, which is a method that functions to evaluate clusters.

The K-Means calculation results in the best group in the iteration experiment of k = 5 with a value of 0.663 which divides the data into 5 clusters: Cluster 0: 165 items, Cluster 1: 133 items, Cluster 2: 197 items, Cluster 3: 165 items, and Cluster 4: 78 items.

CONCLUSION

From the results of tests carried out on rapidminer, 5 clusters were obtained based on the type and class of hospitals in West Java, namely Cluster 0 with 165 Items, Cluster 1 with 133 Items, Cluster 2 with 197 Items, Cluster 3 with 165 Items, Cluster 4 with 78 Items. The total of all datasets tested is 738 items. The results of grouping in this final project are expected to be the basis for researchers in conducting clusters based on the type and class of hospitals in West Java. Future researchers are expected to be able to add supporting attributes in knowing the type and class.

Author's declaration

Authors' contributions and responsibilities

The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation and discussion of results. The authors read and approved the final manuscript.

Funding

Write down the research funding, if any.

Availability of data and materials

All data are available from the authors.

Competing interests

The authors declare no competing interest.

REFERENCES

- Abdullah, D., & Saputra, D. (2016). Aplikasi Penjadwalan Pengadaan Barang Menggunakan Algoritma Apriori. *Ojs.iuli.Ac.Id*, January. <https://ojs.iuli.ac.id/index.php/eng/article/download/29/27>
- Hermanto, T. I., & Muhyidin, Y. (2020). Analisis Data Sebaran Bandwidth Menggunakan Algoritma DbSCAN Untuk Menentukan Tingkat Kebutuhan Bandwidth Di Kabupaten Purwakarta. *Rabit: Jurnal Teknologi Dan Sistem Informasi Univrab*, 5(2), 130–137. <https://doi.org/10.36341/rabit.v5i2.1388>
- Khatisma, C. (2014). Klasifikasi Kelompok Konsumen Menggunakan Metode K-Means Dan Segmentasi Model Fuzzy Rfm (Recency, Frequency, Monetary). 1–11. <https://repository.uin-suska.ac.id/3547/>
- Machmud, R. (2008). Manajemen Mutu Pelayanan Kesehatan. *Jurnal Kesehatan Masyarakat Andalas*, 2(2), 186. <https://doi.org/10.24893/jkma.2.2.186-190.2008>
- Nisak, U. K. (2020). Buku Ajar Statistik Di Fasilitas Pelayanan Kesehatan. In *Buku Ajar Statistik Di Fasilitas Pelayanan Kesehatan*. <https://doi.org/10.21070/2020/978-623-6833-94-0>
- RI, K. K. (2022). Standar Akreditasi Rumah Sakit. *Syria Studies*, 7(1), 37–72. https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civilwars_12December2010.pdf%0Ahttps://thin.k-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625
- Sano, A. V. D., & Nindito, H. (2016). Application of K-means algorithm for cluster analysis on poverty of provinces in Indonesia. *ComTech: Computer, Mathematics and Engineering Applications*, 7(2), 141–150.
- Song, J. Y., Choi, J. Y., Lee, J. S., Bae, I. G., Kim, Y. K., Sohn, J. W., ... & Cheong, H. J. (2013). Clinical and economic burden of invasive pneumococcal disease in adults: a multicenter hospital-based study. *BMC infectious diseases*, 13(1), 1–9.
- STEI INDONESIA. (2017). Bab iii metoda penelitian 3.1. Bab III Metoda Penelitian, Bab iii me, 1–9.
- Suhartina, I. (2019). Analisis Kuantitatif Ketidaklengkapan Pengisian Resume Medis Berdasarkan Program Quality Assurance (Suatu Studi di Rumah Sakit Delta Surya Sidoarjo). *Jurnal Kesehatan Vokasional*, 4(2), 80. <https://doi.org/10.22146/jkesvo.43948>
- Supartiningsih, S. (2017). Kualitas Pelayanan Kepuasan Pasien Rumah Sakit. *Jurnal Medicoeticolegal Dan Manajemen Rumah Sakit*, 6 (1): 9–15, Januari 2017, 6(1), 9–14. <https://doi.org/10.18196/jmmr.6122.Kualitas>
- Yuli Mardi. (2019). Data Mining: Klasifikasi Menggunakan Algoritma C4.5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). *Jurnal Edik Informatika. Jurnal Edik Informatika*, 2(2), 213–219.