Jurnal Elektronik Ilmu Komputer Udayana
Volume 8, No 2. November 2019

p-ISSN: 2301-5373
e-ISSN: 2654-5101

Implementation of the CBR (Case-Based Reasoning) Method in Cases of Caesarean section

Ni Wayan Wiantari [a1], I Wayan Supriana [a2]

[a]Informatics Department, Udayana University
Bali, Indonesia
[1]wayanwiantari11@gmail.com
[2]iwayansupriana@gmail.com

Abstract

CBR (Case-Based Reasoning) method is a reasoning method that uses old knowledge to overcome new problems. CBR will provide solutions to new cases by looking at old cases that are closest to new cases. One case that can use the CBR method is a case of the Caesaran section because several factors affect the Caesaran section as well as features in the system, including, age, number of pregnancies, time of delivery, blood pressure, and heart status so that not everyone can do surgery, Cesar. In this study, a system was used to determine whether a patient could have a Caesaran section or not by using the CBR method and calculate similarity using Naive Bayes. The percentage correlation value of each feature is sought using SPSS because each feature has a different effect on the results. The number of Caesaran section data was 80 data, in this study were divided into 70% training data (56) and 30% testing data (24). Where the new case data will be compared with the old case data in the database, and then the similarity criteria are calculated based on the existing formula. The results of testing of 24 data testing there are 5 data whose results are incompatible and 19 data whose results are by the data before it is shared. So that the accuracy of the Caesaran section with the CBR method using Naive Bayes is 79%.

Keywords: Caesarean section, CBR, Naive Bayes

1. Introduction

Caesaran section is the process of giving birth to the fetus, placenta, and membranes through the abdominal wall by making incisions in the abdominal wall and uterus. Caesaran section can be done if the mother is unable to give birth through a normal process. The operation is carried out with the aim that the safety of the mother and baby can be handled properly. Therefore many patients believe that giving birth by Caesaran section will be better for the mother and baby than the normal childbirth process. However, this operation still has some risks, especially in women with a history of Caesaran section in the previous birthing process [1].

According to WHO the reasonable rate of labor by Caesaran section is 5-10% of all births. It turns out that throughout the world the rate of Caesaran section is increasing rapidly, while the numbers in Indonesia are not yet known with certainty. However, 1 out of every 10 women who give birth every year has had a Caesaran section [2]. On the other hand death or illness that occurs in a Caesarean section, depends on the factors that underlie the operation. Factors affecting the Caesaran section include age, number of pregnancies, time of delivery, blood pressure, and heart status [2].

One method that can be used in cases of Caesaran section is the CBR method. Case- based reasoning (CBR) is a method of reasoning using an approach that focuses on problem- solving based on knowledge from previous cases. In detail, the CBR is divided into four stages, including: retrieve, which is to find a case that is most similar to a new case that will be evaluated, then reuse

information or knowledge that has been stored on a case basis to solve new problems, revise is to improve the proposed solution, and retain to store knowledge that will later be used to solve problems on an existing case base [4].

In studies using the CBR method on the problems of pests and diseases of cassava plants by determining the similarity weights of 5, 3, 2, taking into account the similarity of new cases to old cases, calculation of the level of accuracy using the KNN method with an accuracy of 67.65% [3]. The Case-Based Reasoning (CBR) method is used in the Toddler Growth application by using the Nearest Neighbor calculation, where new case data will be compared with the old case data in the database, and then the similarity criteria are calculated based on the formula or applicable provisions. The results of the database can produce output data in the form of a comparison between old cases and new cases for determining the similarity criteria for under-five growth cases, which are influenced by gender, age, weight, and height, but for the accuracy of the application of CBR on this issue has not been explained [4] In predicting the age of birth using Naive Bayes explained that Naive Bayes is a classification method with a formula that is simple and easy to apply and the Naive Bayes method has a fairly high degree of accuracy compared to other methods [5].

Given the importance of the value of the results of a patient's diagnosis and therapy to be stored because this is very useful for patients undergoing treatment or when seeing a doctor in the future, it is necessary to make a knowledge-based system using the Case-Based Reasoning (CBR) method. this is using a Caesaran section, with user output that can be performed Caesaran section or not using the CBR method that is looking at the old case base and calculating similarity using Naive Bayes. Caesaran section has several factors that influence the results, of all the features, tested the correlation between features and results in SPSS so that the percentage of correlation values also affect the results of identification.

2. Research Methods

2.1 Data Collection

In this study using secondary data obtained through the UCI Machine Learning Repository Center for Machine Learning and Intelligent Systems. A large section dataset is collected from the "Application of Decision Tree Data Mining Algorithm in Health Operations" obtained from UCI data. This dataset consists of 80 cases, consisting of five attributes which are the most important characteristics of the problem. Following features:

Table 1. Features of Caesaran section

Number	Feature	Information
1	Age	In the form of numbers
2	Number of pregnancies	In the form of numbers
3	Time of delivery	Early, on time, late
4	Blood pressure	Low, medium, high
5	Heart status	Yes, No

From each feature, the correlation value in SPSS is tested to find out which features have the biggest correlation with the results. The correlation value of each of these features will be used in the calculation of similarity, namely the calculation of a new case approach on an old case basis.

V6	Pearson Correlation	.078	.145	-.166	-.036	.353**	1
	Sig. (2-tailed)	.492	.200	.141	.753	.001	
	N	80	80	80	80	80	80

Figure 1. SPSS Results for Each Caesaran section

From the correlation value of each feature the percentage is calculated, as follows: Number of correlations = 0.778

Age : (0.078/0.778) * 100% = 10%
Number of pregnancies : (0.145/0.778) * 100% = 19%
Time of delivery : (0.166/0.778) * 100% = 21%
Blood pressure : (0.036/0.778) * 100% = 5%
Heart status : (0.353/0.778) * 100% = 45%

2.2 Case Indexing

In applying the CBR method an indexing process is needed at the retrieving stage, which is taking the problem or case that is most similar to the case base. In applying the CBR method an indexing process is needed at the retrieving stage, which is taking the problem or case that is most similar to the case base.

Table 2. Case Indexing

Number	Age	Number of pregnancies	Time of delivery	Blood pressure	Heart status	Class
1	30	1	0	1	0	0
2	32	1	0	2	1	1
3	26	2	1	1	1	0
4	25	1	0	0	0	0
5	40	1	0	1	1	1
6	32	2	0	2	1	1
7	27	2	0	1	1	1
8	26	2	2	1	0	1
9	28	3	0	2	0	1
10	33	1	1	1	0	0
11	31	2	2	1	0	0

In the indexing process that has been carried out above, as for the explanation of each label in the table above, it is as follows:

a. Age : Age of the patient
b. Number of pregnancies : The number of pregnancies the patient has experienced
c. Time of delivery : 0 (early), 1 (on time), 2 (late)
d. Blood pressure : 0 (low), 1 (medium), 2 (high)
e. Heart status : 0 (has heart disease), 1 (has no disease heart)
f. Class : 0 (can be a Caesaran section), 1 (can't be a Caesaran section)

Decision Making Flowchart

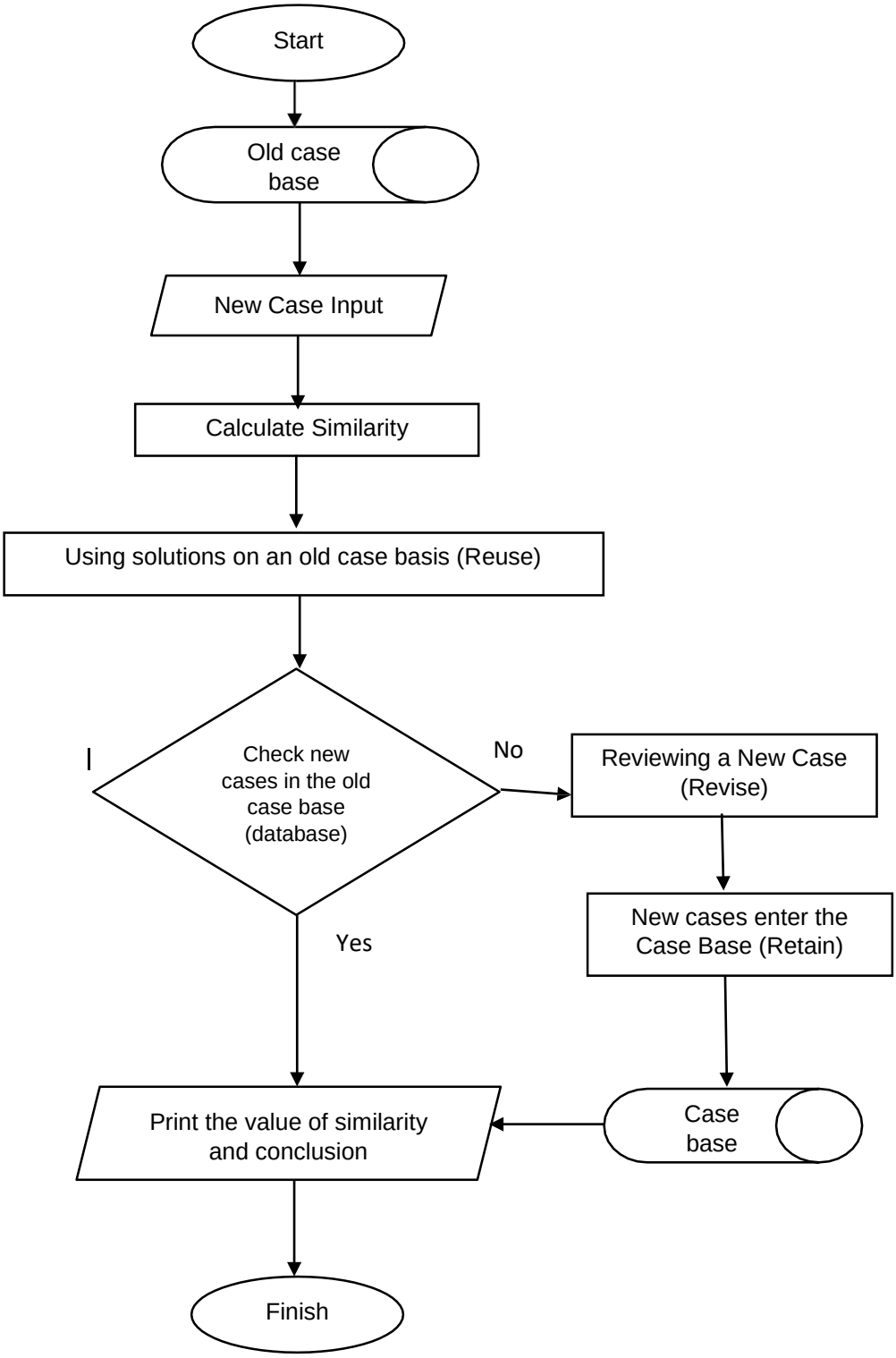


Figure 2. System Flowchart

The picture above is a flow chart of the system created. The system that was made has been incorporated into an old case base. The user enters a new case, after which it enters the

retrieve process by calculating the similarity between the new case and the old case using naive Bayes. Then enter the reuse process by using the solution based on the old case, this happens when the new case with the old case is very similar or the same. Enter the test to check whether new cases already exist based on old cases, if there are then print the similarity value and the user's conclusion can be Caesaran or not. If not, then enter the revision process, namely reviewing new cases based on existing cases in the database, then entering the retain process, namely new cases entering the case base. So that the new case was the basis of the case. Finally, print the similarity value and the user's conclusion can be Caesaran or not.

2.3 Inference Engine

The process in Case-Based Reasoning (CBR) is a match between a new case entered by the user and the case base that is in the database. Then the system calculates the value of similarity using Naive Bayes. By calculating the level of similarity of a case against a new case based on some attributes that are defined based on a certain weighting and then the level of similarity of the whole attribute or feature. The advantage of using Naive Bayes classification is that this method only requires a little training data to make what is needed for classification [5]

Information:

$$P(M_i \mid S) = \frac{P(S \cap M_i)}{P(S)} \quad \text{.....(1)}$$

P (Mi | S) = The probability that each class or feature is calculated S = total cases calculated according to the old case basis
M = number of classes or features counted

The process in Case-Based Reasoning (CBR) is a match between a new case entered by the user and the case base that is in the database. The process of Case-Based Reasoning made is as follows:

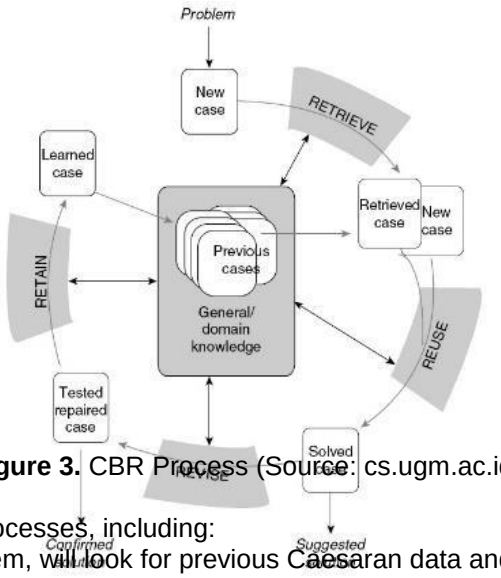


Figure 3. CBR Process (Source: cs.ugm.ac.id)

In the Caesaran system has 4 CBR processes, including:

a. Retrieve, which is a Caesaran system, will look for previous Caesaran data and assess the similarity or similarity to the new Caesaran data.
b. Reuse is to reuse a problem or case to try to solve the problem or case and the Caesaran system will make adjustments to the condition of the old case to the current condition.

c. Revise, Caesaran surgery system will provide conclusions based on the most similar and comparable Caesaran section data.
d. Retain, if a Caesaran section matches the case, the system will store it and use it for future case data

3. Result and Discussion

3.1 System Implementation

The basis of Case-Based Reasoning is solving problems using information stored in old cases. In the expert system that has been made in this study, using training data as a base of old cases and testing data as testing for new cases. The Caesaran system that has been made is as follows:

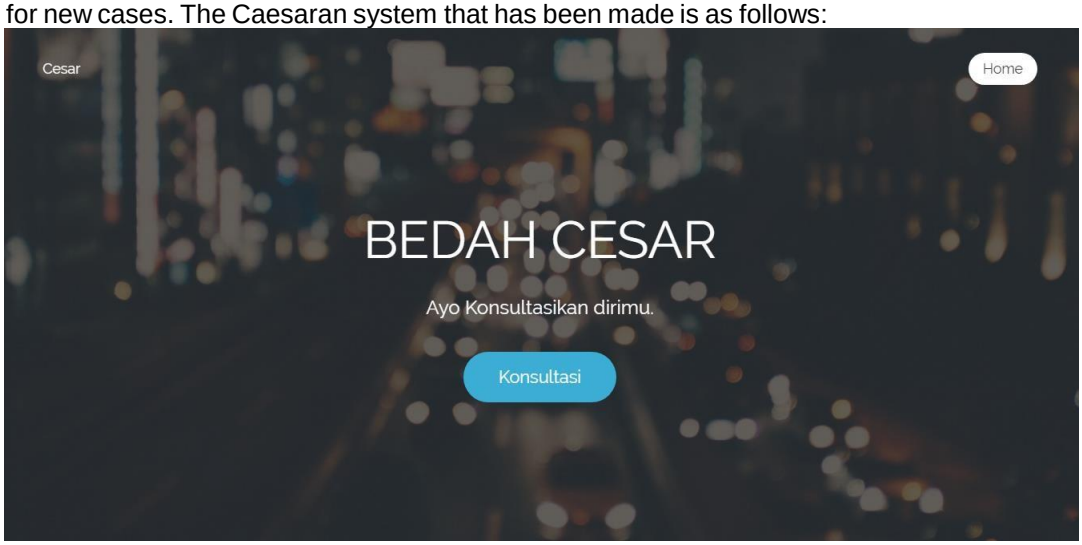


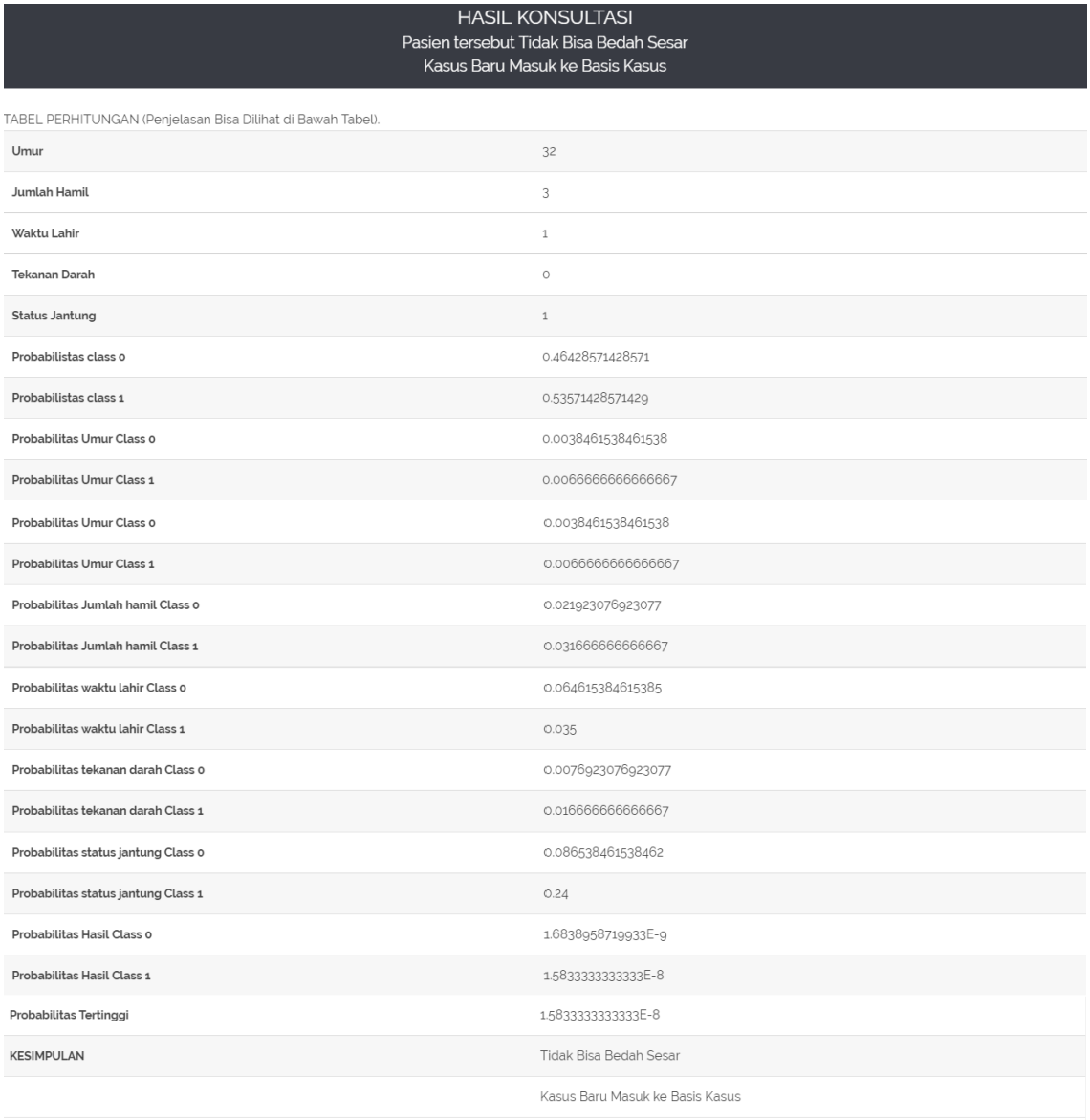
Figure 4. Initial Display of the System

In the picture above is the initial appearance or dashboard of a Caesaran system that has been made. If the user wants to do a consultation, the user can press the consultation button in blue. When the user presses the consultation button it will be directed to the consultation page, as shown below.

Figure 5. Display Data Input

In the picture above is the user display for inputting data. Consists of age, number of pregnancy, time of birth, blood pressure, and heart status. When the user has finished filling all the data, the user presses the blue submit button at the bottom of the table. The submit button will direct the user to the results of the consultation, like the image below.

p-ISSN: 2301-5373
e-ISSN: 2654-5101



HASIL KONSULTASI
Pasien tersebut Tidak Bisa Bedah Sesar
Kasus Baru Masuk ke Basis Kasus

TABEL PERHITUNGAN (Penjelasan Bisa Dilihat di Bawah Tabel).

Umur	32
Jumlah Hamil	3
Waktu Lahir	1
Tekanan Darah	0
Status Jantung	1
Probabilistas class 0	0.46428571428571
Probabilistas class 1	0.53571428571429
Probabilitas Umur Class 0	0.0038461538461538
Probabilitas Umur Class 1	0.0066666666666667
Probabilitas Umur Class 0	0.0038461538461538
Probabilitas Umur Class 1	0.0066666666666667
Probabilitas Jumlah hamil Class 0	0.02192307692307
Probabilitas Jumlah hamil Class 1	0.031666666666667
Probabilitas waktu lahir Class 0	0.064615384615385
Probabilitas waktu lahir Class 1	0.035
Probabilitas tekanan darah Class 0	0.007692307692307
Probabilitas tekanan darah Class 1	0.016666666666667
Probabilitas status jantung Class 0	0.086538461538462
Probabilitas status jantung Class 1	0.24
Probabilitas Hasil Class 0	1.6838958719933E-9
Probabilitas Hasil Class 1	1.5833333333333E-8
Probabilitas Tertinggi	1.5833333333333E-8
KESIMPULAN	Tidak Bisa Bedah Sesar
	Kasus Baru Masuk ke Basis Kasus

Figure 6. Consultation Results

The picture above is the result of consultation when the user input age = 32, the number of pregnancy = 3, time of birth = on time, blood pressure = low, and heart status = do not have heart disease. Then the similarity calculation of each class and feature is displayed, the biggest class 1 conclusion is that the user cannot have a Caesaran section, and is checked whether a new case is in the case base, if there is none then it is displayed as the conclusion above, new cases enter case base, so the new case becomes a case base on the database.

3.2 Application of the CBR Method

1. Retrieve Process

In the retrive process, the probability of new cases is calculated with old cases using Nayve Bayes. The new case taken is a new case as tested in the above system. The following is the manual completion of the calculation of the new case:

Class Probability

P(Result = 0 = Caesar) = 26/56 = 0.464286 P(Results = 1 =
No Caesar) = 30/56 = 0.535714 Probability of Each Feature
P(Age= 32 | Result =0) = 1/26 * 10% = 0.003846
P(Age = 32 | Result =1) = 2/30* 10% = 0.006666667
P(Number of pregnancies = 3 | Result =0) = 3/26* 19% = 0.021923077 P(Number of pregnancies = 3 | Result =1) = 5/30 * 19% = 0.031666667 P(Time of delivery = 1 (on time) | Result =0) = 8/26 *21% = 0.064615385 P(Time of delivery = 1 (on time) | Result =1) = 5/30 *21% = 0.035 P(Blood pressure = 0 (low) | Result = 0) = 4/26 *5% = 0.007692308 P(Blood pressure = 0 (low) | Result = 1) = 10/30 *5% = 0.016666667 P(Heart status = No = 1 | Result = 0) = 5/26 *45% = 0.086538462 P(Heart status = No = 1 | Result = 1) = 16/30 *45% = 0.24

Multiply the probability of the class and each feature Result =0 (Can Caesaran Section)

0.464286 * 0.003846 * 0.021923077 * 0.064615385 * 0.007692308 * 0.086538462

$= 1.68383E^{-9}$

Result=1 (Can't Ceserean Section)

0.535714 * 0.006666667 * 0.031666667 * 0.035 * 0.016666667 * 0.24

$= 1.583333E^{-8}$

The biggest value is 1.583333E-8 which cannot be a Caesaran. Then according to the calculations in the system above with manual calculations.

2. Reuse Process

In the reuse process, a check is made of the similarity between new cases and the existing case base in the database. Especially in the case features and also the results of the case.

3. Revise Process

In the revision process, checking new cases is based on existing case bases in the database. If a new case already exists then a revision is not carried out, but if some data or features are different or not the same then a revise is done with the highest probability result.

4. Retain Process

If a new case does not exist on a case or database basis, then that case becomes the base case with the class that has the highest probability.

3.3 Accuracy Results

Researchers have already tested Caesaran systems. There were 80 cases of Caesaran section, researchers divided into 2 data, 70% trending data (56) and 30% testing data (24). A total of 24 testing data have been tested on the system, found 19 by the results and 5 do not match, from the test found that the accuracy of the Caesaran system was 79%.

4. Conclusions

CBR (Case-Based Reasoning) method is a reasoning method that uses old knowledge to overcome new problems by calculating similarity using Naive Bayes. In this study, the new case data will be compared with the old case data in the database, then the similarity criteria will be

calculated based on the existing formula. The percentage correlation value of each feature is sought using SPSS because each feature has a different effect on the results. The data used in this study is Caesaran case data. The number of Caesaran section data was 80 data, in this study were divided into training data as much as 70% (56) and testing data; as much as 30% (24). In the test results obtained by the accuracy of the Caesaran section with the CBR method using Naive Bayes is 79.

References

[1] Agus. 2014. Hubungan antara Peningkatan Angka Persalinan Seksio Caesar dengan Program Jampersal di RSUD Moewardi Surakarta. Fakultas Ilmu Kesehatah Universitas Muhammadiyah Surakarta.

[2] Amin. 2018. “UCI Machine Learning Repository. [Online] https://archive.ics.uci.edu/ml/datasets/Caesarian+Section+Classification+Dataset

[3] Minarni, dkk. 2017. Case-Based Reasoning (Cbr) Pada Sistem Pakar Identifikasi Hama Dan Penyakit Tanaman Singkong Dalam Usaha Meningkatkan Produktivitas Tanaman Pangan

[4] Mukhammad, dkk. Sistem Pakar Pertumbuhan Balita Berbasis Web Dengan Metode Case Based Reasoning

[5] Naisha, Yogiek. 2018. Aplikasi Prediksi Usia Kelahiran Dengan Metode Naive Bayes. Jurnal SIMETRIS