Scientific Journal of Informatics
Vol. 7, No. 2, November 2020

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Decision Making System to Determine Childbirth Process with Naïve Bayes and Forward Chaining Methods

Putri Laksita Kumalasari[1], Riza Arifudin[2], Alamsyah[3]

[1,2,3]Computer Science Department, Faculty of Mathematics and Natural Sciences, State University of Semarang, Indonesia
Email: [1]putrilaksitak@students.unnes.ac.id, [2]rizaarifudin@mail.unnes.ac.id, [3]alamsyah@mail.unnes.ac.id

Abstract

Childbirth is the last stage before the infant comes into the world. There may be incidents that could cause death in the process of childbirth for mothers and infants. Lack of knowledge and attention to the labor process can increase maternal mortality rate. Maternal mortality rate in Indonesia was recorded at 190 per 100,000 live births on 2015. The figure is still far from the fifth Millennium Development Goals target of 102 per 100,000. The increasing development of technology in health informatics to provide health care more effective can be used to help overcome the problems of pregnant women. To reduce maternal mortality rates, a web-based expert system is perfect one for use. Naïve Bayes method is a simple, fast and high accuracy method. Forward Chaining method is a inference method that performs a fact or statement that starts from the condition (IF) then to the conclusion (THEN). Based on analysis of the method obtained with 233 patients data on childbirth process using expert system, the Naïve Bayes method has accuracy in diagnosing by 90.99% while Forward Chaining method accuracy is 86.70%.

Keywords: Childbirth, Naïve Bayes, Forward Chaining

1. INTRODUCTION

Maternal mortality is an indicator of determining the degree of public health. Maternal mortality describes the risk of obstetrics occurring in each pregnancy calculated by the number of all mothers dying in a given year per 100,000 live births at the same time. The maternal mortality rate in Indonesia ranges from 190 per 100,000 live births according to the Indonesian Demographic and Health survey on 2015[1]. The figure is still far from the global target of MDGs (Millennium Development Goals) in 2015 which is 102 per 100,000 live births and Sustainable Development Goals (SDGs) of 70 per 10,000 of live births to the period 2030[2].

A web-based expert system can be used to assist people in determining the choice of several alternatives given to a decision. Clinical predictions are the result of data processing using data mining methods. So it can predict the occurrence of risk that can occur in pregnant women. The reduced risk of occurrence in pregnant women can affect the number of maternal mortality and infant mortality rate.

Naïve Bayes algorithm is a classification method using probability and statistical methods. Bayes Theorem provides a simple rule to calculate conditional probabilities. Disadvantages of Naïve Bayes method is based on three factors, namely training data noise, bias, and variance [3]. Forward chaining method is a Forward chaining method start from the input information (IF) first and then to the conclusion (THEN) with IF-THEN rule [4]. Forward Chaining method is a search method by tracking existing information and then combining rules to generate a conclusion or a goal.

The previous research [5] conducts research by comparing the level of accuracy between Naïve Bayes, Decision Tree and Random Forest Algorithm. In the study, Naïve Bayes method has the highest accuracy of 83.43%. A research [6] using naïve Bayesian method to identification of Tuberculosis with 237 data sample and the results given by the system is 85.95% accuracy. This paper [7] conducts research to determination type lenses glasses using forward chaining with 75.7% accuracy. A research [8] combining two methods of Forward Chaining and Naïve Bayes method with the accuracy results given by the system is 96.05%.

This study aims to compare Naïve Bayes and Forward Chaining methods in predicting the childbirth process. We use Naïve Bayes and Forward Chaining methods to determine the more accurate and precise algorithm.

2. METHODS

2.1 Naïve Bayes

Bayes theorem algorithm is an algorithm that assumes all independent or interdependent attributes given by values in class variables [9]. Naïve Bayes uses a collection of sample training data that has been labeled to estimated model parameters[10]. Naïve Bayes method is a classifications with probability and statistical methods. The probability of a specific feature in the data appears as a member in the set of probabilities and is derived by calculating the frequency of each feature value within a class of a training data set. The training dataset is a subset, used to train a classifier algorithm by using known values to predict future, unknown values[11]. Equation (1) are commonly used on Naïve Bayes[12].

$$P(Y|X) = \frac{P(X|H)P(H)}{P(X)} \tag{1}$$

In this case:

X = data with an unknown class

H = data hypothesis X is a specific class

$P(H|X)$ = H hypothesis probability based on condition X (posterior probability)

$P(H)$ = probability of hypothesis H (prior probability)
$P(X|H)$ = probability of X based on the conditions in hypothesis H
$P(X)$ = probability of X

2.2 Forward Chaining

Forward Chaining method is a method of the inference engine to start reasoning or tracking data form facts that lead to a conclusion [13]. Forward Chaining method is also called data-driven [14]. The main advantages from Forward Chaining method are these methods work well when the problem stems from collecting or bringing together information and looking for conclusions from such information. The other side Forward Chaining has the possibility cannot identify some more important facts from other facts and cannot give conclusions from the problem[6]. The Forward Chaining process can be modeled in Figure 1.

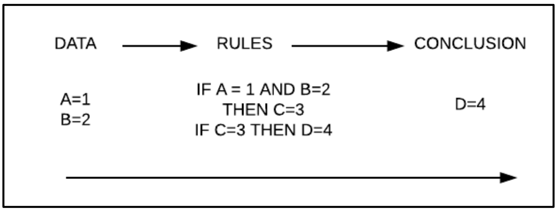

Figure 1. Process Forward Chaining

The process of Forward Chaining method as in Figure 1 they are 1) the inference engine will collect data or facts, data or facts will be used to determine the conclusion of the problem, 2) the rules made are used to find a conclusion, then make a knowledge base in Forward Chaining, the inference engine will looking for rules in the knowledge base accordance with these data or facts, the system checks the rule again looking for new matches, rules that were previously fired are ignored, this procedure continues until no match is found 3) from the rules chosen, it will obtained a conclusion.

3. RESULTS AND DISCUSSION

3.1 Data Processing

3.1.1 Naïve Bayes

In this research conducted 3 steps in completing the calculation Naïve Bayes i.e. the prior probability, likelihood probability, posterior probability[15]. The prior calculations are performed to compare many members of a class with the entire sample data. The likelihood calculation is a calculation of the probability values of each attribute to its class, which allows raising classes when an attribute is selected. And the final step is the posterior calculation of the calculations used to draw conclusions. In the posterior calculation will be a comparison to the posterior value of each existing class. The highest posterior value is selected as the classification result.

a. Prior Probability

A process compares many members of a class with the entire sample data. Prior calculations can be done with the Equation (2).

$$P = \frac{X}{A} \tag{2}$$

In this case:

P = prior value
X = total data per class
A = total data

Here are the conditions of sample 1 patients used to do the calculation of Naive Bayes.

1. Mother's age: 25 years old
2. Caesarean Surgical History: None
3. The location of the baby breech: No
4. Cephalopelvic Disporpotion (CPD): No
5. Placenta previa: No
6. Severe Preeclampsia: No
7. Oligohidroamnion: No
8. Hypertension: Yes

Based on Example 1 prior probability on Naïve Bayes using Equation(2) where:

$P(R01) = \frac{120}{233} = 0.515$
$P(R02) = \frac{113}{233} = 0.485$

b. Likelihood Probability

Llikelihood probability is a calculation of the probability values of each attribute to its class, possibly a class when an attribute is selected. Likelihood probability can be seen in Equation(3).

$$L = \frac{F}{B} \tag{3}$$

In this case:

L = Likelihood value
F = Number of feature data per class
B = Amount of all data per class

From the previous calculation, we know that P(R01) is 0.515 and P(R02) is 0.485. The next step is to find the value of *L* or likelihood. Using Equation (3) we can find likelihood probability shown as follows:

$P(G01 = NO|R01) = \frac{102}{184} = 0.554$
$P(G01 = NO|R02) = \frac{82}{184} = 0.446$

Explanation about likelihood value of sample 1 is shown in Table 1.

Table 1. Likelihood value of sample 1

No	Maternity code	Conditions							
		G01= No	G02= No	G03= No	G04= No	G05= No	G06= No	G07= No	G08= No
1.	R01	0.554	0.625	0.581	0.566	0.524	0.543	0.553	0
2.	R02	0.446	0.375	0.419	0.434	0.476	0.457	0.446	1

c. Posterior Probability

Posterior probability is a result of likelihood probability in the form of an attribute probability to the class used to look for opportunities for the inclusion of certain characteristic samples in a class. In this process acquired a final probability to take a conclusion. The posterior process of calculation can be seen in Equation 4.

$$P = (H|E) = P(H) \times P(E|H) \quad (4)$$

The classification results are done by comparing the posterior value of existing classes. The highest posterior value is chosen as the classification result. Based on Example 1 then the posterior probability of the Naïve Bayes using Equation (4) is as follows:

Posterior P1=0.554*0.625*0.581*0.566*0.524*0.543*0.553*0=0
Posterior P2=0.446*0.375*0.419*0.434*0.476*0.457*0.446*1=0.002951

The highest posterior value is P2 so the childbirth process recommendation in Example 1 is posterior P2(Sectio Caesarea).

3.1.2 Forward Chaining

In this study the transfer of knowledge from experts into the system has been described into two categories namely maternity solutions and mother's condition. Delivery solution can be seen in Table 2 and childbirth symptoms can be seen in Table 3.

Table 2. Maternity solutions

No	Maternity Solutions	Code
1.	Normal	R01
2.	Sectio Caesarea	R02

Table 3. Mothers condition

No	Conditions	Code
1.	Age	G01
2.	Caesarean Surgical History	G02
3.	The location of the baby breech	G03

4.	Cephalopelvic Disproportion(CPD)	G04
5.	Placenta Previa	G05
6.	Severe Preeclampsia	G06
7.	Oligohidroamnion	G07
8.	Hypertension	G08

a. Rule Base

Rule base or production rules are used to diagnose forward chaining method. Rule base is obtained from the results of a decision table that serves to store the data of labor and mother conditions that are arranged by the relationship of each attribute. The rule base table is shown in Table 4.

Table 4. Rule base forward chaining

Rule	**IF**	**THEN**
1	G01=NO^G02=NO^ G03=NO^G04=NO^ G05=NO^G06=NO^ G07=NO^G08=NO	R01

According Table 3 can be concluded that the patient who has a condition not in accordance with R01 then the patient will be advised to perform childbirth in Sectio Caesarea or R02. The decision table produced in Table 3 is used as a reference in drafting production rules.

b. Decision making on Forward Chaining

In the process of taking a conclusion on the forward chaining method there are two possibilities that can occur ie the selection of conclusions detected or inferences undetected. The selection of conclusions is detected if the end result of the delivery process indicates one type of conclusion, if the final result is more than one conclusion then the search is stopped and no conclusions are matched. From Sample 1 the search process with forward chaining method can be written as follows:

G01=No
G02=No
G03=No
G04=No
G05=No
G06=No
G07=No
G08=Yes

The search process with Forward Chaining method is shown in Table 5.

Table 5. Table of forward chaining search process

No	Conditions Code	Solution Detected	Directions Searches	Description
1.	G01=No	R01,R02	G02	Search continues
2.	G02=No	R01,R02	G03	Search continues
3.	G03=No	R01,R02	G04	Search continues

4.	G04=No	R01,R02	G05	Search continues
5.	G05=No	R01,R02	G06	Search continues
6.	G06=No	R01,R02	G07	Search continues
7.	G07=No	R01,R02	G08	Search continues
8.	G08=Yes	R02	Finished	Solution detected

Based on Table 5 it is known that the result of the calculation using Forward Chaining method is R02(Sectio Caesarea).

3.2 Testing

Based on 233 data that has been tested using the expert system in each method, there are 212 corresponding data using Naive Bayes method and 21 uncorresponding data and there are 202 corresponding data and 31 uncorresponding data with Forward Chaining method. Figure 2 displays the results of system using Naïve Bayes and Forward Chaining methods.

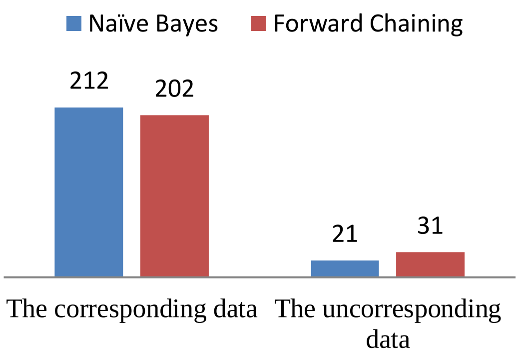

Figure 2. System diagnosis result

After the calculation process using Naïve Bayes and Forward Chaining method with 233 patients, then the result will be process to get the value of accuracy. The accuracy value can be calculated using Equation (5).

$$\text{The corresponding accuracy value} = \frac{\textit{The amount of data is accurate}}{\textit{total data}} \times 100 \qquad (5)$$

The accuracy value produced by the system using Equation (5).

1. Naïve Bayes

The amount of data	= 233
The corresponding data	= 212
The uncorresponding data	= 21
The corresponding accuracy value	$= \frac{212}{233} \times 100\% = 90.99\%$

2. Forward Chaining

The amount of data	= 233

The corresponding data = 202

The uncorresponding data = 31

The corresponding accuracy value $= \frac{202}{233} \times 100\% = 86.70\%$

3. CONCLUSION

Based on explanation, Naïve Bayes method is a simple, fast method and has a higher level of accuracy. While Forward Chaining method is a method of piping so as to get a conclusion must go through the appropriate flow based on experts. If the flow is done according to the rule base then it will get the result of the conclusion, otherwise if there is one plot that does not match then the case will not find the conclusion. The result of system accuracy by implementing 233 data into the system, Naïve Bayes method has an accuracy rate of 90.99% and Forward Chaining method has an accuracy rate of 86.70%.

4. REFERENCE

[1] Dewi, A., Bekti, N. K., & Supriyatiningsih. (2019). Maternal Mortality Evaluation: A Case Study in Bantul, Yogyakarta. *Journal of Maternal and Child Health*, 342-350.

[2] Alkema, L., Chou, D., Hogan, D., Zhang, S., Moller, A. B., Gemmill, A., & Say, L. (2016). Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *The Lancet*, *387*(10017), 462-474. Mukherjee, S., & Sharma, N. (2012). Intrusion Detection using Naive Bayes Classifier with Feature Reduction. *Procedia Technology*, *4*, 119.

[3] Windriyani, P., Wiharto, W., & Sihwi, S. W. (2013). Expert System for Detecting Mental Disorder with Forward Chaining Method. *International Conference on ICT for Smart Society*, 1-7.

[4] Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm. *Procedia Computer Science*, 161, 765-772.

[5] Agustin Trihartati, S., & Adi, C. K. (2016). An Identification of Tuberculosis (TB) Disease in Humans using Naïve Bayesian Method. *Scientific Journal of Informatics*, *3*(2), 99.

[6] Pramesti, A. A., Arifudin, R., & Sugiharti, E. (2016). Expert System for Determination of Type Lenses Glasses using Forward Chaining Method. *Scientific Journal of Informatics*, 177-188.

[7] Perbawawati, A. A., Sugiharti, E., & Muslim, M. A. (2019). Bayes Theorem and Forward Chaining Method On Expert System for Determine Hypercholesterolemia Drugs. *Scientific Journal of Informatics*, 116.

[8] Patil, T. R., & Sherekar, M. S. (2013). Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification. *International Journal of Computer Science and Applications*, 256-261.

[9] Azalia, F. Y., Bijaksana, M. A., & Huda, A. F. (2019). Name indexing in Indonesian translation of hadith using named entity recognition with naïve Bayes classifier. *Procedia Computer Science*, 157, 142-149.
[10] Dimitoglou, G., Adams, J. A., & Jim, C. M. (2012). Comparison of the C4.5 and a Naive Bayes Classifier for the Prediction of Lung Cancer Survivability. *Journal of Computing*, 1-9.
[11] Wibowo, A., Manongga, D., & Purnomo, H. D. (2020). The Utilizatin of Naive Bayes and C.45 in Predicting the Timeliness of Students' Graduation. *Scientific Journal of Informatics*, 99-112.
[12] Gusman, A. P., & Hendri, H. (2019). Expert system to diagnose child development growth disorders with forward chaining method. *Journal of Physics: Cnference Series*, 1-8.
[13] Al-Ajlan, A. (2015). The comparison between forward and backward chaining. *International Journal of Machine Learning and Computing*, 5(2), 106-113.
[14] Insani, M. I., Alamsyah, A., & Putra, A. T. (2018). Implementation of Expert System for Diabetes Diseases using Naïve Bayes and Certainty Factor Methods. *Scientific Journal of Informatics*, 5(2), 185-193.