International Journal of Information System & Technology
Akreditasi No. 158/E/KPT/2021 | Vol. 6, No. 1, (2022), pp. 93-100

Application of Naive Bayes Method For Diagnosis of Pregnancy Disease

Embun Fajar Wati[1], *Budi Sudrajat*[2]

[1,2]*Universitas Bina Sarana Informatika*

[1]*embun.efw@bsi.ac.id*, [2]*budi.bst@bsi.ac.id*

Abstract

The risk of pregnancy can be known by early detection of pregnancy with risk factors, so that health workers can know more about treatment. In diagnosing a disease in the field of medicine requires tools such as the application of artificial intelligence, one of which is an expert system. One method that can be applied in expert systems is naive bayes. In this study, naive bayes for the process of diagnosing the disease during pregnancy was done by including symptoms that appear in pregnant women. The stage of research is the collection of data from previous research journal articles with the same theme, but different methods and other journal articles with the same theme and different from the naive bayes method. The next stage is data analysis with naive bayes calculations of patient symptoms and validation, namely comparing the results of naive bayes calculations with expert calculations. The results obtained were 14 patients out of 20 patients, which is 70% have the same results between experts with calculations with naive bayes. The results showed that the calculation of symptoms with naive bayes was sufficient to give valid and feasible results to use.

Keywords: *pregnancy, naïve bayes, symptoms, expert system, patient*

1. Introduction

Information technology can be used in all fields such as offices, business, hospitality, education and health [1] Health is related to pregnancy. Pregnancy is a process that is very desirable by some married women. During pregnancy, the mother and fetus are a unit of function that cannot be separated [2]. There are often pregnancy disorders because women usually look healthy as if nothing happened, but the condition does not mean there are no problems in the fetus. Many women also ignore their symptoms in the process of pregnancy resulting in death [3]. Lack of knowledge and information about pregnancy diseases can cause delays in pregnant women to know the presence of diseases in pregnancy [2].

The risk of pregnancy can be known by early detection of pregnancy with risk factors so that health workers can know more about treatment. In addition to maintaining fetal growth and health, pregnancy care is needed to prevent complications and death during childbirth. To avoid or overcome the risk, it is necessary to do regular examinations at midwives or doctors [4]. Midwives are one of the health care places for pregnant women. The advantages are the close distance, and the affordable cost. Many queues of pregnant women who do pregnancy examinations, thus making midwives and doctors speed up the examination, so that the solution provided is less than opimal [5].

In diagnosing a disease in the field of medicine requires tools such as the application of artificial intelligence, Artificial intelligence is one of the fields of computer science that utilizes computers so that they can behave intelligently like humans [6]. Expert systems are one example of artificial intelligence [7]. The expert system is a computer application program that seeks to imitate the reasoning process of an expert in solving, specification problems or can be said to be a clone of an expert because the science is stored in the knowledge base to process problem solving [8]. When the user operates the computer, the

ISSN : 2580-7250

expert system will issue some factual questions to get conclusions and explanations [9]. The application of expert systems is seen as a way to implement the knowledge of experts in a particular field in a program so that they can make decisions and make intelligent reasoning [10]. The limited number of medical personnel, can be helped by the existence of an application of expert systems, without intending to replace experts. With the help of expertise, information is summarized in the database as a source of handling disease diagnoses until the solution will be done as a step to solve the problem [11]. So that the public can know the first action that must be done before going to the doctor for further treatment [12]. The application of expert systems has become a commonly applied thing, especially in the field of medicine [13].

One method that can be applied in expert systems is naive bayes which is one of the methods that can be used to classify data [14]. Naive Bayes algorithm predicts future opportunities based on previous experience [15]. The Naive Bayes Classifier method is also a simple probabilistic classification that calculates a set of probabilities by summing the frequencies and combinations of values from a given dataset [16]. This method was chosen because it is easy to apply to work independently, namely a feature in a data is not related to the presence or absence of other features in the same data. This method is considered good enough to determine the probability in determining the result [17]. Naive Bayes often works much better in most complex real-world situations than expected [18].

Researchers previously researched pregnancy diagnosis using forward chaining methods. While in the current study, researchers use naive bayes for the process of diagnosing the disease during pregnancy is done by including symptoms that appear in pregnant women. Through these symptoms will be done calculations to obtain a probability value on each type of disease. Types of diseases that have a high final probability value will be taken as a result of the diagnosis of the expert system.

2. Research Methodology

The research methodology used in this research includes the process of data collection and data analysis.

2.1. Data Collection

Data collection is done by searching for journal articles online. The data collected is disease data (table 1) as many as 5 diseases, complaint data (table 2) as many as 32 complaints or symptoms, and patient data and perceived complaints (table 3).

Table 1. Disease Data

Code	Disease
P01	Hyperemesis third trimester
P02	Hyperemesis gravidarum in level 1
P03	Hyperemesis gravidarum in level 2
P04	Hyperemesis gravidarum in level 3
P05	Mild preeclampsia
P06	Preeclampsia
P07	Eclampsia

Table 2. Symptom Data

Code	Symptom
G01	Bleeding in young and old pregnant
G02	Amniotic water comes out before its time
G03	Excessive nausea or vomiting
G04	Upper abdominal pain
G05	Dry and dirty tongue
G06	Dehydration
G07	Decreased appetite
G08	Weight Loss

Code	Symptom
G09	Sunken eyes
G10	Increased pulse rate & lower blood pressure
G11	Pulse frequency about 100 beats/minute
G12	Looks weak and limp (not fit)
G13	Yellow eyes
G14	Difficulty defecating
G15	Decreased skin elasticity
G16	The faster the pulse frequency above 100 beats/ minute
G17	Small pulse as blood volume drops
G18	Increased body heat or fever
G19	Urine slightly until it does not come out urine
G20	Vomiting and mixed blood
G21	Decreased awareness
G22	Out of sight
G23	Headache or dizziness
G24	Excess weight or greater weight gain
G25	Swollen face or other parts of the body
G26	Blood pressure between 140/90 to 160/110
G27	Proteinuria +1
G28	Proteinuria +2
G29	Proteinuria +3 or more
G30	Tensions over 160/110
G31	Heart a pain
G32	Seizures

Explanation :
P : Disease
G : Symptom

Table 3. Patient and Symptom Data

Patient	Age	Symptom
A	26	G24
B	20	G06, G23, G24, G25, G26, G31
C	30	G01, G02
D	35	G23, G25, G26, G27
E	37	G03, G04, G05, G06, G07, G08, G09, G13, G14, G15
F	28	G23, G24, G25
G	27	G04, G23
H	22	G23, G24, G28, G30
I	27	G04, G23, G24, G25, G26
J	37	G03, G04, G05, G07, G10, G11, G24, G26
K	28	G04, G10, G13, G14, G15, G17, G20, G21, G22, G24
L	24	G01, G03, G24, G26, G27
M	39	G24
N	21	G02, G04, G07, G14, G23, G24, G25, G26
O	23	G01, G02, G03
P	29	G24, G25, G28, G30
Q	26	G25, G26, G27, G30
R	29	G03, G07, G13, G23, G24, G27, G28, G29
S	36	G04
T	30	G23, G24, G26, G30

2.2. Data Analysis

After the data are collected, the next stage is to analyze the data by grouping the diseases and symptoms complained of (table 4). Then the calculation is done by the naive bayes method and validation is done, which is to compare the results of the naive bayes calculation with the results of the expert diagnosis.

Table 4. Disease dan Symptom Data

Disease	Symptom
P01	G01, G02
P02	G03, G04, G05, G06, G07, G08, G09, G10, G11, G12
P03	G03, G04, G05, G08, G09, G13, G14, G15, G16, G17, G18, G19, G20
P04	G03, G04, G05, G10, G13, G14, G17, G18, G19, G20, G21, G22
P05	G23, G24, G25, G26, G27
P06	G13, G23, G24, G25, G28, G30
P07	G13, G21, G22, G23, G24, G25, G29, G30, G31, G32

Here's a calculation of the probability of naive bayes (covid implementasi) :
P(A|B)=(P(B|A)P(A)) / (P(B)
Explanation :
P (A|B) : Chance A if it knows the condition of type of symptom B
P(B|A) : Chance of evidence B if hypothesis A is known
P(B) : Probability B without looking at any evidence
P(A) : Chance of evidence of symptoms A

3. Result and Discussion

Result and discussion discuss the stages of research methods, namely the data analysis of data, calculations using the naive bayes method.
Examples of calculations for patient O:

3.1. Specifies the nc value for each class

The first stage is to find the value of nc for each disease based on the symptoms that appear.

a) P1 : Hyperemesis third trimester
   n : 1
   m : 32
   p : 1/7 = 0,14
   G01 : 1
   G02 : 1
   G03 : 0
b) P2 : Hyperemesis gravidarum in level 1
   n : 1
   m : 32
   p : 1/7 = 0,14
   G01 : 0
   G02 : 0
   G03 : 1
c) P3 : Hyperemesis gravidarum in level 2
   n : 1
   m : 32
   p : 1/7 = 0,14
   G01 : 0
   G02 : 0
   G03 : 1
d) P4 : Hyperemesis gravidarum in level 3
   n : 1
   m : 32
   p : 1/7 = 0,14
   G01 : 0
   G02 : 0
   G03 : 1

e) P5 : Mild preeclampsia
n : 1
m : 32
p : 1/7 = 0,14
G01 : 0
G02 : 0
G03 : 0

f) P6 : Preeclampsia
n : 1
m : 32
p : 1/7 = 0,14
G01 : 0
G02 : 0
G03 : 0

g) P7 : Eclampsia
n : 1
m : 32
p : 1/7 = 0,14
G01 : 0
G02 : 0
G03 : 0

3.2. Calculate the value of P(A|B) and calculate the value P (B)

The second stage is to calculate the probability value for each disease based on symptoms.

a) P1 : Hyperemesis third trimester (HTT)
P(1|HTT) = (1+32.0,14)/1+32 = 0,17
P(2|HTT) = (1+32.0,14)/1+32 = 0,17
P(3|HTT) = (0+32.0,14)/1+32 = 0,14
P(HTT) = 0,14

b) P2 : Hyperemesis gravidarum in level 1 (HGIL1)
P(1|HGIL1) = (0+32.0,14)/1+32 = 0,14
P(2|HGIL1) = (0+32.0,14)/1+32 = 0,14
P(3|HGIL1) = (1+32.0,14)/1+32 = 0,17
P(HGIL1) = 0,14

c) P3 : Hyperemesis gravidarum in level 2 (HGIL2)
P(1|HGIL2) = (0+32.0,14)/1+32 = 0,14
P(2|HGIL2) = (0+32.0,14)/1+32 = 0,14
P(3|HGIL2) = (1+32.0,14)/1+32 = 0,17
P(HGIL2) = 0,14

d) P4 : Hyperemesis gravidarum in level 3 (HGIL3)
P(1|HGIL3) = (0+32.0,14)/1+32 = 0,14
P(2|HGIL3) = (0+32.0,14)/1+32 = 0,14
P(3|HGIL3) = (1+32.0,14)/1+32 = 0,17
P(HGIL3) = 0,14

e) P5 : Mild preeclampsia (MP)
P(1|MP) = (0+32.0,14)/1+32 = 0,14
P(2|MP) = (0+32.0,14)/1+32 = 0,14
P(3|MP) = (0+32.0,14)/1+32 = 0,14
P(MP) = 0,14

f) P6 : Preeclampsia (PL)
P(1|PL) = (0+32.0,14)/1+32 = 0,14
P(2|PL) = (0+32.0,14)/1+32 = 0,14

P(3|PL) = (0+32.0,14)/1+32 = 0,14
P(PL) = 0,14

g) P(1) = 0,14P7 : Eclampsia (EL)
P(1|EL) = (0+32.0,14)/1+32 = 0,14
P(2|EL) = (0+32.0,14)/1+32 = 0,14
P(3|EL) = (0+32.0,14)/1+32 = 0,14
P(EL) = 0,14

3.3. Counting P(A|B) x P(B) for each B

The third stage is to multiply the probability value of each disease by each symptom.

a) P1 : Hyperemesis third trimester (HTT)
P(HTT) x (P(1|HTT) x P(2|HTT)) x P(3|HTT))
0,14 x 0,17 x 0,17 x 0,14 = 0,000566

b) P2 : Hyperemesis gravidarum in level 1 (HGIL1)
P(HGIL1) x (P(1| HGIL1) x P(2| HGIL1)) x P(3| HGIL1))
0,14 x 0,17 x 0,17 x 0,17 = 0,000688

c) P3 : Hyperemesis gravidarum in level 2 (HGIL2)
P(HGIL2) x (P(1| HGIL2) x P(2| HGIL2)) x P(3| HGIL2))
0,14 x 0,14 x 0,14 x 0,17 = 0,000466

d) P4 : Hyperemesis gravidarum in level 3 (HGIL3)
P(HGIL3) x (P(1| HGIL3) x P(2| HGIL3)) x P(3| HGIL3))
0,14 x 0,14 x 0,14 x 0,17 = 0,000466

e) P5 : Mild preeclampsia (MP)
P(MP) x (P(1| MP) x P(2| MP)) x P(3| MP))
0,14 x 0,14 x 0,14 x 0,14 = 0,000384

f) P6 : Preeclampsia (PL)
P(PL) x (P(1| PL) x P(2| PL)) x P(3| PL))
0,14 x 0,14 x 0,14 x 0,14 = 0,000384

g) P(1) = 0,14P7 : Eclampsia (EL)
P(EL) x (P(1| EL) x P(2| EL)) x P(3| EL))
0,14 x 0,14 x 0,14 x 0,14 = 0,000384

From the results of the calculation above, the highest multiplication value obtained by P2 is hyperemesis gravidarum disease in level 1 with a value of 0.000688.

3.4. Validation

Expert validation is the matching of the results issued by the system regarding diseases, according to the expert's knowledge base. The validation results of the calculation of the symptoms of all patients can be seen in table 5 below.

Table 5. Expert Validation

Patient	Naïve Bayes	Expert
A	Preeclampsia	Preeclampsia
B	Mild preeclampsia	Mild preeclampsia
C	Hyperemesis third trimester	Hyperemesis third trimester
D	Eclampsia	Mild preeclampsia
E	Eclampsia	Hyperemesis gravidarum in level 1
F	Preeclampsia	Preeclampsia
G	Eclampsia	Eclampsia
H	Eclampsia	Eclampsia
I	Hyperemesis gravidarum in level 3	Preeclampsia
J	Hyperemesis gravidarum In Level 1	Hyperemesis gravidarum in level 1
K	Eclampsia	Hyperemesis gravidarum in level 1
L	Preeklampsia	Preeklampsia

Patient	Naïve Bayes	Expert
M	Eclampsia	Eclampsia
N	Mild preeclampsia	Mild preeclampsia
O	Hyperemesis gravidarum In Level 1	Hyperemesis gravidarum in level 1
P	Eclampsia	Preeclampsia
Q	Preeclampsia	Preeclampsia
R	Preeclampsia	Preeclampsia
S	Hyperemesis gravidarum In Level 1	Hyperemesis gravidarum in level 1
T	Eclampsia	Mild preeclampsia

The validation results showed that 14 patients out of 20 patients, namely 70% had the same results between experts with calculations with naive bayes. The results showed that the calculation of symptoms with naive bayes was sufficient to give valid and feasible results to use.

4. Conclusion

The conclusion that can result from the calculation of symptoms in pregnant women is that the naive bayes method is feasible to use for the diagnosis of symptoms of disorders in pregnancy with results showing that 70% i.e. 14 out of 20 pregnant women patients have valid results with an expert diagnosis. For further research will be done by comparison of several other methods that are expected to increase the percentage of the validity of hadil diagnosis between calculations and experts.

References

[1] P. A. D. Putra, I. K. A. Purnawan, and D. P. S. Putri, "Sistem Pakar Diagnosa Penyakit Mata dengan Fuzzy Logic dan Naïve Bayes," *MERPATI*, vol. 6, no. 1, pp. 35–46, 2018.

[2] H. Effendi, D. Ariyadi, and I. Sabroto, "Sistem Pakar Diagnosa Penyakit pada Ibu Hamil," *TEKNOMATIKA*, vol. 10, no. 1, pp. 9–20, 2020.

[3] Ekojono, B. S. Andoko, and K. A. Romadlan, "Sistem Pakar untuk Mendiagnosa Gangguan Kehamilan dengan Metode Naive Bayes," in *Seminar Informatika Aplikatif Polinema (SIAP)*, 2021, pp. 161–166.

[4] Q. Hasanah, A. Andrianto, and M. A. Hidayat, "Sistem Informasi Posyandu Ibu Hamil dengan Penerapan Klasifikasi Resiko Kehamilan Menggunakan Metode Naïve Bayes," *Berk. SAINSTEK*, vol. 6, no. 1, pp. 1–9, 2018.

[5] E. F. Wati and A. Puspitasari, "Expert System for Diagnosing Pregnancy Complaints by Forward Chaining," *Sink. J. dan Penelit. Tek. Inform.*, vol. 5, no. 1, pp. 7–16, 2020.

[6] R. Restari, S. Sinurat, and Suginam, "Rancangan Aplikasi Sistem Pakar Diagnosa Penyakit Mononukleosis dengan Metode Naive Bayes," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 3, pp. 403–408, 2020.

[7] R. Rizky, Susilawati, Z. Hakim, and L. Sujai, "Sistem Pakar Deteksi Penyakit Hipertensi Dan Upaya Pencegahannya Menggunakan Metode Naive Bayes Pada RSUD Pandeglang Banten," *JUTIS*, vol. 7, no. 2, pp. 138–144, 2019.

[8] E. F. Wati, L. Hakim, and A. P. Sari, "Sistem Pakar Diagnosa Penyakit Mata Pada Manusia Dengan Metode Forward Chaining," *JISAMAR*, vol. 2, no. 4, pp. 29–38, 2018.

[9] M. Nanda and A. Mubarok, "Sistem Pakar Penentuan Penyakit Polip Hidung dengan Rinosinusitis Menggunakan Naïve Bayes Berbasis Android," *Infortech*, vol. 2, no. 2, pp. 159–165, 2020.

[10] A. S. Puspaningrum, E. R. Susanto, and A. Sucipto, "Penerapan Metode Forward Chaining Untuk Mendiagnosa Penyakit Tanaman Sawi," *Informatics J.*, vol. 5, no. 3, pp. 113–120, 2020.

[11] Yuliyana and A. S. R. M. Sinaga, "Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Naive Bayes," *Fountain Informatics J.*, vol. 4, no. 1, pp. 19–23, 2019.
[12] F. Ramadhana, Fauziah, and Winarsih, "Aplikasi Sistem Pakar Untuk Mendiagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Berbasis Website," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 4, no. 3, pp. 320–329, 2020.
[13] W. Setiawan and S. Ratnasari, "Sistem Pakar Diagnosis Penyakit Mata Menggunakan Naïvebayes Classifier," in *Seminar Nasional Sains dan Teknologi*, 2014, pp. 1–6.
[14] R. I. Borman and M. Wati, "Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes," *J. Ilm. Fak. Ilmu Komput.*, vol. 9, no. 1, pp. 25–34, 2020.
[15] Y. B. Widodo, S. A. Anggraeini, and T. Sutabri, "Perancangan Sistem Pakar Diagnosis Penyakit Diabetes Berbasis Web Menggunakan Algoritma Naive Bayes," *J. Teknlogi Inform. dan Komput. MH. Thamrin*, vol. 7, no. 1, pp. 112–123, 2021.
[16] I. Gunaawan and Y. Fernando, "Sistem Pakar Diagnosa Penyakit Kulit Pada Kucing Menggunakan Metode Naive Bayes Berbasis Web," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 239–247, 2021.
[17] M. R. Handoko and Neneng, "Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 50–58, 2021.
[18] D. Simanjuntak and A. Sindar, "Sistem Pakar Deteksi Gizi Buruk Balita Dengan Metode Naïve Bayes Classifier," *J. Inkofar*, vol. 1, no. 1, pp. 54–60, 2019.