BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 1, Mare 2024 , Hal. 31 - 36
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i1.1210
https://journal.fkpt.org/index.php/BIT

Designing an Expert System Application to Diagnose Dengue Fever Patients Using the Naïve Bayes Method

Fransisca Angelia Sebayang [1]*
[1] Program Studi Sistem Informasi, STMIK Kristen Neumann Indonesia, Medan
Email: [1,*] fransiscasebayang@gmail.com
(* : coressponding author: fransiscasebayang@gmail.com)

Abstrak- Demam berdarah adalah penyakit menular yang disebabkan oleh virus dengue dan ditularkan melalui gigitan nyamuk Aedes aegepty dan Aedes albopictus. Naive Bayes adalah pengklasifikasi probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari kumpulan data yang diberikan. Pada penelitian ini akan dilakukan analisis data penyakit demam berdarah dengue dengan menggunakan metode naive bayes dengan menggunakan 9 variabel. Aplikasi sistem pakar untuk mendiagnosa penyakit demam berdarah ini dibuat dengan aplikasi berbasis web, sehingga dapat diakses oleh masyarakat secara luas, selain itu aplikasi ini juga dapat membantu petugas medis untuk mengambil keputusan dalam mendiagnosa penyakit demam berdarah. Dari hasil penelitian, penggunaan metode Naïve Bayes pada aplikasi ini dikarenakan hasil probabilitas nilai akurasi dari metode Naïve Bayes yang mendekati akurasi dari para ahli. Pada penelitian ini dilakukan 3 kali percobaan dengan menggunakan 4 data uji dan hasilnya menunjukkan bahwa berdasarkan ketiga percobaan tersebut, nilai akurasi berkisar antara 96% hingga 97% yang menandakan bahwa ketepatan klasifikasi memiliki nilai error yang kecil. Sehingga klasifikasi menggunakan Naïve Bayes pada penelitian ini dapat diterapkan untuk menentukan klasifikasi kejadian demam berdarah dengue.

Kata Kunci: Naïve Bayes; Sistem Pakar; Pasien Demam Berdarah

Abstract− Dengue fever is an infectious disease caused by the dengue virus and transmitted through the bites of Aedes aegepty and Aedes albopictus mosquitoes. Naive Bayes is a simple probabilistic classifier that calculates a set of probabilities by summing the frequency and combination of values from a given dataset. In this study, we will analyze dengue hemorrhagic fever disease data using the naive bayes method using 9 variables. This expert system application to diagnose dengue disease is made with a web-based application, so that it can be accessed by the public at large, besides that this application can also help medics to make decisions in diagnosing dengue disease. From the results of the study, the use of the Naïve Bayes method in this application is due to the results of the probability of the accuracy value of the Naïve Bayes method which is close to the accuracy of the experts. In this study, 3 experiments were conducted using 4 test data and the results showed that based on the three experiments, the accuracy value ranged from 96% to 97%, indicating that the classification accuracy had a small error value. So that classification using Naïve Bayes in this study can be applied to determine the classification of dengue hemorrhagic fever incidence.

Keywords: Naïve Bayes; Expert System; Dengue Fever Patients

1. PENDAHULUAN

Health is the most important thing for humans because everyone can experience health problems. Therefore, we must be able to maintain our health so that we are not attacked by diseases, one of which is dengue fever. An unbalanced body, a disturbed metabolic system, and decreased immunity are some of the causes of health problems [1]. Everyone's immune system is different, so not everyone has symptoms of the disease. As a result of the disease attacking the body, all tasks are stopped [2].

Most people usually do not pay attention to their health, especially dengue fever, which has some symptoms similar to ordinary fever, therefore, do not want to see a doctor. Due to cost constraints, lack of medical personnel, less effective services to patients, and limited working time of doctors [3]. Dengue fever is an infectious disease caused by the dengue virus and transmitted through the bites of Aedes aegepty and Aedes albopictus mosquitoes [4]. Dengue fever is one of the infectious diseases that often causes outbreaks and causes death [3].

So the need for a technology that is able to adopt the way humans think, namely artificial intelligence technology. An expert system is a part of artificial intelligence where the system tries to adopt human knowledge to computers so that computers can solve problems as commonly done by experts [5] [6]. With an expert system, users can interact with computers to solve certain problems. The implementation of expert systems can be applied in the health sector in addition to information for the public, especially people with certain diseases to find out the initial diagnosis, also as a tool for doctors to make decisions [7] [8].

Based on this, a system is needed to identify the type of disease accompanied by fever based on other symptoms that appear. So that it can help in the process of diagnosis and further treatment. The system in question is a disease diagnosis system accompanied by fever using the Naive Bayes Classifier method [9] [10]. This application system for diagnosing diseases accompanied by fever is based on an Expert System or expert system with an application display in the form of a website [11].

Naive Bayes Classifier is a simple probability classifier based on Bayes' theorem. The advantage of the Naive Bayes Classifier is that it only requires a small amount of training data to estimate parameters (means and variances of variables) in the classification process. In the process, the Naïve Bayes Classifier assumes that the presence or absence of a feature in a class is not related to the presence or absence of other features in the same class [2] [12].

Therefore, a system is needed to identify the type of disease accompanied by fever based on additional symptoms so that it can help in the diagnosis and treatment process. The system in question is an application system for diagnosing diseases

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 1, Mare 2024 , Hal. 31 - 36
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i1.1210
https://journal.fkpt.org/index.php/BIT

accompanied by fever using the Naive Bayes Classifier method. This system is based on an expert system, or expert system, and has an application display through the website [13] [14]. This expert system application will start the workflow by filling in the patient's personal data and selecting the symptoms complained of. Then, the diagnosis results page is displayed to view the diagnosis results. It is hoped that this.

2. METODOLOGI PENELITIAN

2.1 Research Stages

This framework is the steps that will be taken in order to solve the problem to be discussed. Figure 2. below is the framework (framework) used in this study.

Figure1. Research Procedure Framework

Based on the framework in Figure 1, each step can be described as follows:

1. Describe the Problem
   Describe the problem clearly to get the results of a comparative study of the naïve Bayes classifier and decision tree methods in the classification of payment type data.
2. Problem Analysis
   The problem analysis step is a step to understand the problem whose scope or boundaries have been determined. By analyzing the predetermined problem, it is hoped that the problem can be well understood.
3. Setting Goals
   Based on the understanding of the problems of the problems, the objectives to be achieved in this study were determined. This goal determines the targets to be achieved, especially those that can overcome existing problems.
4. Studying Literature Related to the Title
   To achieve the goal, then studied some of the literature that is expected to be used. Then the literature studied is selected which will be used in this study. Literature sources were obtained from books and journals.
5. Data Collection
   The data needed is data that will be used as material for research, namely data on types of payment deals from Kaggle.
6. System Analysis
   System analysis is quite important to do, because here the author must know the weaknesses of the system, obstacles, constraints and opportunities that are not able to be achieved by the current system in order to find alternative solutions to the problem.
7. System Design
   The system is designed using rapid miner tools to manage payment type data by looking at the accuracy and percentage of calculations for each method.
8. Program Structure
   Program Structure Design is a design that describes the relationship between a communication system with other communication systems.
9. Analysis Results
   At this stage it will provide the results of the research analysis of the naïve Bayes classifier and decision tree models which produce the percentage model of each method.

2.2 Expert System

Expert System or Expert System is also known as Knowledge Based System, which is a computer application intended to help make decisions or solve problems in a specific field [16]. This system works by using knowledge and analysis methods that have been defined in advance by experts who are in accordance with their fields of expertise. [17].

2.3 Naïve Bayes Classifier

Naive Bayes models are a group of extremely fast and simple classification algorithms that are often suitable for very high-dimensional datasets. Because they are so fast and have so few tunable parameters, they end up being very useful as a quick-and-dirty baseline for a classification problem [18]. The Naïve Bayes classifier is a fairly simple probabilistic

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 1, Mare 2024 , Hal. 31 - 36
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i1.1210
https://journal.fkpt.org/index.php/BIT

classification method. The calculation carried out in this method is on a set of opportunities by adding up the frequency and combination of values from the dataset owned. The Naïve Bayes Classifier method assumes that each variable in each category is independent of the other [19]. Naive Bayes is one classic machine learning algorithm based on a Bayesian network, which is usually applied to classification problems and has excellent performance. When using naive Bayes for classification, from the perspective of probability, for the item to be classified, it needs to calculate the probability of the item appearing in each target category, then choose the most significant probability and find the corresponding category, as the classification result [20].

The Naïve Bayes Classifier algorithm has no relationship between one attribute and another, or in other words one attribute has no effect on other attributes, even though these attributes may be related [21]. Classification capabilities in Naïve Bayes are based on the Bayes theorem. In addition, the Naïve Bayes Classifier can also be said to be a method for making predictions in the future based on previous experience [19]. Based the Bayes theorem formula can be written as follows:

$P(H|X) = P(X|H).P(H)$
$P(X)$ ..... (1)

where:

X : unknown class data
H : hypothesis from data X, whose class specifications are known
$P(H|X)$: the probability of conjecture H based on state X (posterior probability)
$P(H)$: the probability of conjecture H (prior probability)
$P(X|H)$: the probability of X based on the state of the conjecture H
$P(X)$: the probability of X

2.4 Dengue Fever Patients

Dengue virus causes dengue fever (DHF). DHF is a disease that is troubling the community to this day. Children are more susceptible to DHF and an estimated 20 million people are infected each year. DHF has become a disease that is most concerned about control and treatment by the World Health Organization (WHO), so that its spread does not become more widespread and reduce casualties [22].

3. HASIL DAN PEMBAHASAN

Contains the results and implementation of research based on the data to be achieved in the research methodology. Starting from data preparation, model application, to model use. The following will describe the results and implementation of the research.

3.1 Data Preparation

Dengue fever patients who conduct examinations where there are symptoms experienced by patients. Then the existing data is collected into data features that will be used in research. The following is the dbd patient examination data: From the data preparation, the dataset is divided into characteristics or features (symptoms), then also added with symptoms of diseases that have similarities with dengue fever, because often we cannot distinguish the type of disease because of the same symptoms as other diseases. The following data is created:

Table 1. Symptom Table

Kode	Gejala	Penyakit		
		DBD	Malaria	Tipes
1	Sakit Kepala	0	1	1
2	Menggigil	0	1	0
3	Demam Tinggi	1	1	1
4	Batuk	1	0	1
5	Mual	1	1	1
6	Pegal Linu	0	1	0
7	Mimisan	1	0	1
8	Anemia	0	1	0
9	Gusi Berdarah	1	0	0
10	Bintik-bintik merah dibadan	1	0	0
11	Bibir berdarah	1	0	0
12	Tubuh Lemas	0	1	1
13	Berkeringat Banyak	0	1	0
14	Nyeri Ulu Hati	1	0	0
15	Bab Bedarah	1	0	0

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 1, Mare 2024 , Hal. 31 - 36
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i1.1210
https://journal.fkpt.org/index.php/BIT

Table 2: Training Data Table

Name	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	Diagnosa
Hafizhan Shidqi	0	0	1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	1	0	0	Positif (1)
Gandhi Wibowo	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	Negatif (0)
Aldio Mahendra Purwandrarto	0	1	0	0	1	1	1	0	1	1	1	1	0	0	1	0	1	0	0	1	Positif (1)
Benny Putra	0	0	1	1	0	1	0	1	0	0	1	0	1	0	0	0	1	1	1	0	Negatif (0)
Vicky Vernando Dasta	1	1	0	1	1	1	0	1	1	1	0	0	1	0	0	1	0	0	1	0	Negatif (0)
Jufianto Henri	0	0	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	1	1	1	Positif (1)
Aan Nuraini	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	1	1	0	0	0	Negatif (0)
Abdur Rahman	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	1	1	1	1	0	Positif (1)
Abdurrahman	0	0	1	1	0	0	1	1	0	0	0	1	1	1	0	0	0	0	1	0	Negatif (0)

Description:
0 = No Symptoms
1 = Has Symptoms
Based on the data above we have labeled positive and negative data based on dengue fever symptoms, we will use this data as training data to apply the naïve bayes model.
Problem solving in detecting diseases caused by mosquito bites with the naïve bayes method.mosquito bites with the naïve bayes method is as follows:
a. First, it is known that the symptoms experienced by stakeholders include: Bleeding including nosebleeds, vomiting blood and bloody stools (G01); high fever suddenly then fever drops and then the fever rises again (G02); decreased appetite (G03).
b. Second, read the data on symptoms and types of mosquitoes. Can be seen in table 1.
c. Third, calculate P(H) based on the probability of predicting the type of mosquito bite.
P(N01) = 6 / 9 = 0.67
P(N02) = 3 / 9 = 0.33
P(N03) = 3 / 9 = 0.33
d. Fourth, calculate the class probability P(X|H).
P(N01|Symptoms) = 1 / 3 = 0.33
P(N02|Symptom) = 2 / 3 = 0.67
P(N03|Symptom) = 2 / 3 = 0.67
P(X|Mosquito Type) = 0.33 * 0.67 * 0.67 = 0.15
e. Fifth, calculate the final probability P(H|X) by multiplying P(H) with P(X|H)
P(N01)P(X|Mosquito Type) = 0.67 * 0.15 = 0.1005
P(N02)P(X|Mosquito Type) = 0.33 * 0.15 = 0.0495
P(N03)P(X|Mosquito Type) = 0.33 * 0.15 = 0.0495
Since the value of 0.1005 is the largest, the symptoms that have been experienced by the stakeholder ar is classified as being bitten by a dengue fever type mosquito.

Figure 2. Web Page of Expert System for Diagnosing Dengue Fever

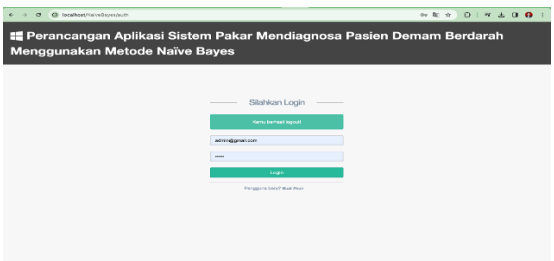

Figure 3. Admin Login Page

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 1, Mare 2024 , Hal. 31 - 36
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i1.1210
https://journal.fkpt.org/index.php/BIT

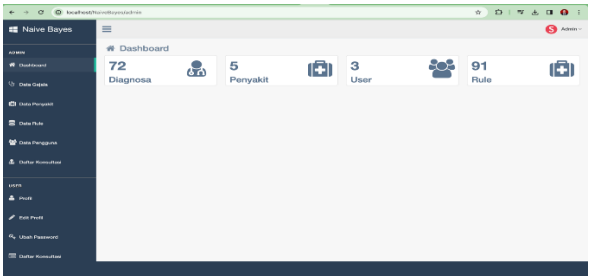

Figure 4. Admin Main Menu Page

4. CONCLUSION

The system will calculate the naïve bayes process with the first step of calculating the mean, standard deviation, and probability using a predetermined formula. Furthermore, the system will calculate the posterior and classification using the naïve bayes method. In testing the naïve bayes method using test data as much as 25 data obtained an accuracy value of 92% with a total positive dengue hemorrhagic fever of 17 people and negative dengue hemorrhagic fever of 8 people. This expert system in the future can be developed based on mobile and also requires more patient data to compare with the results in this study, so that the accuracy level is better.

REFERENCES

[1] A. D. Permana, I. M. A. Suyadnya, and D. C. Khrisne, "the Design of Expert System for Determining the Initial Diagnosis of Tropical Infectious Diseases in Indonesia With Naive Bayes Method-Based Android," *J. Terap. Teknol. Inf.*, vol. 2, no. 2, pp. 123–133, 2019, doi: 10.21460/jutei.2018.22.112.

[2] H. Maradona, M. Rifqi, K. Yasdomi, and V. Desiyanti, "Sistem Pakar Diagnosa Penyakit Demam Berdarah Dengan Metode Naive Bayes Classifier," *RJOCS (Riau J. Comput. Sci.*, vol. 8, no. 2, pp. 109–115, 2022, doi: 10.30606/rjocs.v8i2.1439.

[3] E. Nurlelah, A. Abdilah, and M. A. Ghani, "Implementasi Algoritma Naive Bayes Pada Sistem Pakar Untuk Diagnosis Penyakit Demam Berdarah Dengue Berbasis Website," *J. Speed-Sentra Penelit. Eng. dan Edukasi*, vol. 11, no. 3, pp. 1–8, 2019.

[4] Y. Bismo and G. Harsanto, "Penerapan Metode Certainty Factor Dan Naïve Bayes Untuk Mendiagnosa Penyakit Akibat Gigitan Nyamuk," *Gener. J.*, vol. 4, no. 2, pp. 49–60, 2020, doi: 10.29407/gj.v4i2.14438.

[5] S. T. Wicaksono, "Identifikasi Pasien Demam Berdarah Menggunakan Data Mining Naive Bayes Pada Puskesmas Paduraksa Pemalang," 2021, [Online]. Available: http://repository.unissula.ac.id/id/eprint/24064.

[6] J. Infokum, "Application_of_Data_Mining_in_the_Classi.pdf," vol. 10, no. 2, pp. 803–809, 2022.

[7] N. H. Alfianty and S. Mulyati, "Penerapan Naïve Bayes untuk Klasifikasi Data Penyakit Pada Anak," *Automata*, 2022, [Online]. Available: https://journal.uii.ac.id/AUTOMATA/article/view/21914.

[8] P. Hasan Putra, M. Syahputra Novelan, and M. Rizki, "Analysis K-Nearest Neighbor Method in Classification of Vegetable Quality Based on Color," *J. Appl. Eng. Technol. Sci.*, vol. 3, no. 2, pp. 126–132, 2022.

[9] E. Rantoso and O. Suria, "Rantoso & Suria, Sistem Pakar Diagnosa Penyakit yang Disertai Demam Menggunakan Metode Naive Bayes Classifier Sistem Pakar Diagnosa Penyakit yang Disertai Demam Menggunakan Metode Naive Bayes Classifier Expert System for Diagnosting Disease with Fever Usi," pp. 1–10, 2018.

[10] A. Wahyu Redhani and N. Hidayat, "Implementasi Metode Naïve Bayes untuk Diagnosa Pengidap Demam Berdarah pada Kelurahan Antasan Besar berbasis Web," vol. 5, no. 12, pp. 5320–5328, 2021, [Online]. Available: http://j-ptiik.ub.ac.id.

[11] P. A. Gatto, R. Maulana Awangga, and R. Andarsyah, "Diagnosis Penyakit Demam Berdarah Menggunakan Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1676–1681, 2023, doi: 10.36040/jati.v7i3.6891.

[12] R. A. Setijaningsih, A. Setiawan, Z. Sugiyanto, I. Farida, B. Widjajanto, and I. Rizqa, "Sistem Manajemen Penegakan Diagnosa Penyakit Typus Dengan Metode Naive Bayes," *J. Transform.*, vol. 20, no. 2, p. 70, 2023, doi: 10.26623/transformatika.v20i2.5792.

[13] I. Budiman and Ferdiansyah, "Implementasi Algoritma Naïve Bayes Pada Sistem Pakar Diagnosa Penyakit Tifoid Berbasis Web," *J. Ilm. Betrik*, vol. 13, no. 03, pp. 295–303, 2022.

[14] D. P. B. B and R. Saptono, Ristu Anggrainingsih, "Academic Articles Classification Using Naive Bayes Classifier (NBC) Method," *J. Teknol. dan Inf.*, vol. 7, no. 2, pp. 74–81, 2018.

[15] M. Desmita, "Sistem Pakar Untuk Mendiagnosa Penyakit Meningitis Menggunakan Metode Certainty Factor Berbasis Web," *J. Sist. dan Inform.*, vol. 12, no. 2, pp. 132–139, 2018, [Online]. Available: http://repository.potensi-utama.ac.id/jspui/handle/123456789/1257.

[16] R. Achmad and A. S. Girsang, "Implementation of naive bayes classifier algorithm in classification of civil servants," *J. Phys. Conf. Ser.*, vol. 1485, no. 1, 2020, doi: 10.1088/1742-6596/1485/1/012018.

[17] Y. Yuliyana and A. S. R. M. Sinaga, "Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Naive Bayes," *Fountain Informatics J.*, vol. 4, no. 1, p. 19, 2019, doi: 10.21111/fij.v4i1.3019.

[18] J. P. Tanjung, F. C. Tampubolon, A. W. Panggabean, and M. A. A. Nandrawan, "Customer Classification Using Naive Bayes Classifier With Genetic Algorithm Feature Selection," *Sinkron*, vol. 8, no. 1, pp. 584–589, 2023, doi: 10.33395/sinkron.v8i1.12182.

[19] A. Tanza and D. T. Utari, "Comparison of the Naïve Bayes Classifier and Decision Tree J48 for Credit Classification of Bank Customers," *EKSAKTA J. Sci. Data Anal.*, vol. 3, no. 2, pp. 70–77, 2022, doi: 10.20885/eksakta.vol3.iss2.art2.

[20] W. Guo, G. Wang, C. Wang, and Y. Wang, "Distribution network topology identification based on gradient boosting decision tree and attribute weighted naive Bayes," *Energy Reports*, vol. 9, pp. 727–736, 2023, doi: 10.1016/j.egyr.2023.04.256.

[21] A. Zainal Macfud, A. Pandu Kusuma, and W. Dwi Puspitasari, "Analisis Algoritma Naive Bayes Classifier (NBC) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 1–8, 2023.

[22] Warna, "Implementasi Algoritma Certainty Factor untuk Mendiagnosa Penyakit yang Disertai Demam," vol. IV, pp. 129–137, 2023.