VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5064

EXPERT SYSTEM FOR DISEASE IDENTIFICATION BASED ON HEMATOCHEZIA SYMPTOMS WITH NAÏVE BAYES METHOD

Dasril Aldo[1*]; Alwendi[2]; Adanti Wido Paramadini[3]; Ilwan Syafrinal[4]; Sapta Eka Putra[5]

Teknik Informatika[1], Teknik Biomedis[3]
Institut Teknologi Telkom Purwokerto[1,3]
https://ittelkom-pwt.ac.id/[1,3]
dasril@ittelkom-pwt.ac.id[1*], adanti@ittelkom-pwt.ac.id[3]

Ilmu Komputer[2]
Universitas Graha Nusantara Padangsidimpuan[2]
https://ugn.ac.id/[2]
alwendi60@gmail.com[2]

Teknik Perangkat Lunak[4]
Universitas Universal[4]
http://uvers.ac.id/[4]
ilwansynl@gmail.com[4]

Manajemen Retail[5]
Universitas Tamansiswa Padang[5]
https://www.unitas-pdg.ac.id/[5]
saptaeka54putra@gmail.com[5]

(*) Corresponding Author

Abstract*—Hematochezia is a common clinical symptom in various gastrointestinal diseases, requiring accurate diagnosis for effective treatment. This study aims to develop an expert system for the rapid and precise identification of hematochezia-causing diseases. The expert system is designed to assist patients in efficiently recognizing diseases, minimizing treatment failure risks. It employs the Naïve Bayes method, a data calculation approach involving summing combinations and frequencies of each dataset. The expert system methodology begins with training using a dataset comprising hematochezia symptoms and corresponding disease diagnoses. The dataset is input into a database as training data. Subsequently, it undergoes classification and training stages. Symptom data can then be processed using the Naïve Bayes method. The system's end result displays probability values for each disease based on provided symptoms. This analysis relies on specific symptoms selected by the user, such as Rectal Pain, Hematochezia, Constipation, Fatigue, and Abdominal Cramps. It yields a Hemorrhoids diagnosis with a posterior probability of 0.514738. In testing with 35 sample cases, the expert system exhibited a remarkable accuracy rate of 94.29%. This expert system efficiently supports disease diagnosis based on hematochezia symptoms, aiding in swift and accurate identification.*

Keywords*: expert systems, hematochezia, identification, naïve bayes.*

Intisari*—Hematochezia adalah gejala klinis yang sering muncul pada berbagai penyakit gastrointestinal. Diagnosis yang akurat diperlukan untuk pengobatan yang efektif. Penelitian ini mengembangkan sistem pakar untuk mengidentifikasi cepat dan akurat penyakit yang menyebabkan hematochezia. Sistem pakar ini dirancang untuk membantu pasien dalam mengidentifikasi penyakit dengan cara yang efisien, sehingga potensi kegagalan dalam pengobatan dapat diminimalisir. Sistem pakar ini menerapkan metode Naïve Bayes yang merupakan metode menghitung kumpulan data dengan cara menjumlahkan kombinasi dan frekuensi dari setiap dataset. Adapun tahapan dari metode sistem pakar ini diawali dengan dilatih menggunakan dataset yang berisi gejala hematochezia dan diagnosis penyakit yang mendasarinya, dengan memasukkan dataset ke dalam database sebagai data pelatihan, Selanjutnya dilakukan proses klasifikasi dataset, selanjutnya melakukan traning, kemudian dapat memasukkan data gejala untuk diproses menggunakan metode Naïve Bayes. Hasil akhir sistem ini akan menampilkan nilai probabilitas untuk setiap penyakit berdasarkan gejala yang diberikan. Analisis ini dilakukan berdasarkan gejala-gejala spesifik yang dipilih oleh* *pengguna, seperti Nyeri pada area rektal, Hematochezia, Konstipasi, Kelelahan, dan Kram atau nyeri perut, menghasilkan diagnosis Hemoroid dengan nilai probabilitas Posterior 0.514738. Dalam pengujian menggunakan 35 sampel kasus, sistem pakar menunjukkan tingkat akurasi yang sangat tinggi, yaitu 94.29%. Sistem pakar ini dapat digunakan untuk mendukung proses diagnosa penyakit berdasarkan gejala hematochezia.*

Kata Kunci: *sistem pakar, hematochezia, identifikasi, naïve bayes.*

INTRODUCTION

Currently, the field of informatics holds great significance in the swift progress of computing e [1] data processing [2], and artificial intelligence [3]. In this study, the Naive Bayes method is employed to create an expert system designed to identify diseases based on specific symptoms, focusing on Hematochezia. The objective is to contribute to the progress of computer science and its practical applications.

The limitation of access to health information is a critical issue for many individuals [4], particularly during emergencies marked by symptoms such as Hematochezia, patients frequently experience anxiety and strive for an accurate initial comprehension of the underlying causes of their symptoms. Nonetheless, a considerable number of individuals face challenges in obtaining trustworthy health information. Consequently, the development of a tool becomes imperative to aid individuals with restricted access in precisely and effortlessly recognizing diseases characterized by Hematochezia symptoms, facilitating prompt and appropriate actions.

This research makes a significant contribution by introducing a Naive Bayes-based expert system designed for patient use in accurately discerning diseases exhibiting Hematochezia symptoms. Furthermore, the study advances the field of computer science by incorporating the Naive Bayes method into the identification process of these diseases, potentially fostering positive developments in computer science knowledge and its practical applications in public health.

The principal aim of this investigation is to devise an expert system that aids individuals, particularly patients, in autonomously and expeditiously pinpointing diseases with precision when presented with Hematochezia symptoms. The expected outcome of this study is to pave the way for further advancements in medical expert systems, capable of tackling diverse challenges associated with diagnosing diseases featuring intricate symptoms like Hematochezia. This, in turn, empowers individuals to take timely and appropriate measures to safeguard their health.

The research methodology involves the application of the Naive Bayes method to formulate an expert system tailored for identifying diseases characterized by Hematochezia symptoms. Utilizing symptom data and patients health histories as input, the method employs probabilities to generate the most probable diagnosis. Through in-depth exploration, experimentation, and performance evaluation, we believe that the implementation of this method will positively impact the enhancement of the medical diagnosis process and the future development of expert systems.

The application of the Naive Bayes method in expert systems has proven to be highly accurate in the early identification of diseases. This is demonstrated by Ain et al., with an accuracy result of 96% in diagnosing stroke [5]. Furthermore, Irfansyah et al. achieved an accuracy of 90% in diagnosing Hepatitis [6]. Subsequent research by Budianto et al. also yielded an accuracy of 92% in diagnosing corn plant diseases [7]. Other research by Santiko et al. with an accuracy value of 93.54% focused on diagnosing chronic kidney disease [8]. Another advancement research by Maliha et al. achieved an accuracy result of 98.2% in diagnosing cancer [9].

The application of the Naive Bayes method in expert systems has demonstrated remarkable accuracy in the prompt detection of diseases across diverse scenarios [10]. Numerous investigations by different scholars, encompassing the diagnosis of diseases such as stroke, hepatitis, corn plant diseases, and chronic kidney disease, have resulted in noteworthy levels of precision. This indicates the substantial potential of the Naive Bayes method to augment accuracy in the early identification of diseases. In addressing the preceding challenge, specifically the restricted access of individuals to health information, employing the Naive Bayes method emerges as a viable and efficient solution. Through the integration of this method into expert system development, individuals with limited access can adeptly and autonomously identify diseases manifesting Hematochezia symptoms and overcome informational barriers that might impede medical interventions. Consequently, the utilization of the Naive Bayes method has potential to enhance the precision of early disease identification.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5064

MATERIALS AND METHODS

This research methodology is designed to develop an expert system based on the Naïve Bayes method in improving the accuracy of diagnosis of gastrointestinal diseases that indicate symptoms of hematochezia. Adopting the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework [11], the study combines quantitative data analysis with input from medical experts. The goal is to produce diagnostic solutions that are not only innovative but also practical as shown in figure 1.

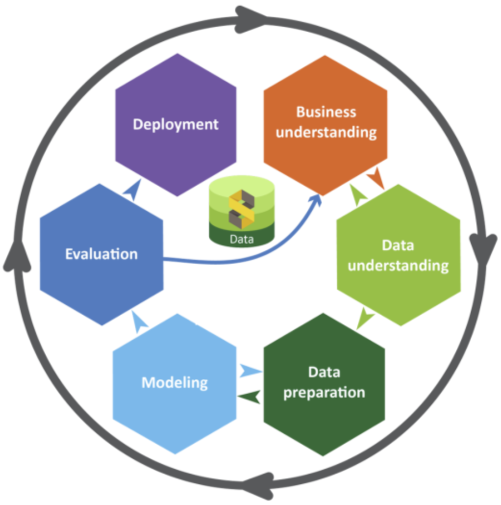

Figure 1. Research Methodology

Figure 1 provides a visual display of the process and steps taken in the study. The details shown in the illustration clearly outline the main phases taken during the study, providing a deeper understanding of the methodology applied.

1. Business Understanding
   This phase underscores the need to develop an expert system that can identify diseases based on symptoms of hematochezia with high accuracy. Because hematochezia is a symptom found in a variety of gastrointestinal diseases, accuracy in diagnosis is critical to effective treatment. The business goal here is to create diagnostic aids that can be used by medical professionals to improve treatment outcomes. The study was designed to fill this gap in the medical field using the Naïve Bayes method, which is expected to predict disease more accurately based on a combination of reported symptoms. Project planning includes determining the necessary resources, schedules, and methodologies.
2. Data Understanding
   At this stage, data were collected from 250 patients with symptoms of hematochezia, including 19 symptoms and 5 different diseases. In addition to quantitative data analysis, interviews with medical experts are conducted to gain a deeper understanding of the associated symptoms and diseases. Medical documentation is also checked to ensure the accuracy of classifying symptoms and diseases. These activities provide important clinical insights that aid in data interpretation and modeling.
3. Data Preparation
   The data is prepared by dividing it into two sets: training (215 patients) and testing (35 patients). Deep data cleansing measures are performed to ensure the consistency and accuracy of the dataset. This includes processes such as imputation for lost data and normalization of data if needed. Next, the data is encoded to satisfy the format required by the Naïve Bayes model, including encoding categorical variables and scaling numerical variables.
4. Modeling
   At the modeling stage, Naïve Bayes was chosen for his ability to classify data based on conditional probabilities. This method is perfect for categorical data such as medical symptoms. The model is trained with a training dataset, which involves adjusting parameters and selecting features. The model is then validated using techniques such as cross-validation to ensure that overfitting does not occur.
5. Evaluation
   After training, the model is evaluated with a test dataset. Metrics such as accuracy, sensitivity, specificity, and predictive value are used to assess model performance. Error analysis is also performed to understand the types of errors made by the model (e.g., false positives and false negatives). This evaluation is important to ensure the reliability and validity of the model in real-world conditions.
6. Application
   The application of this expert system involves the integration of the model into clinical practice. The deployment plan includes creating a friendly user interface to allow medical professionals to enter symptoms and receive a diagnosis. The system will also be equipped with a monitoring module to collect feedback and system performance in real use. The final report will document the entire process, from data collection to evaluation results, as well as recommendations for further development.

To create an accurate expert system for identifying the Based on Hematochezia Symptoms with Naïve Bayes Method, The stages are as follows:

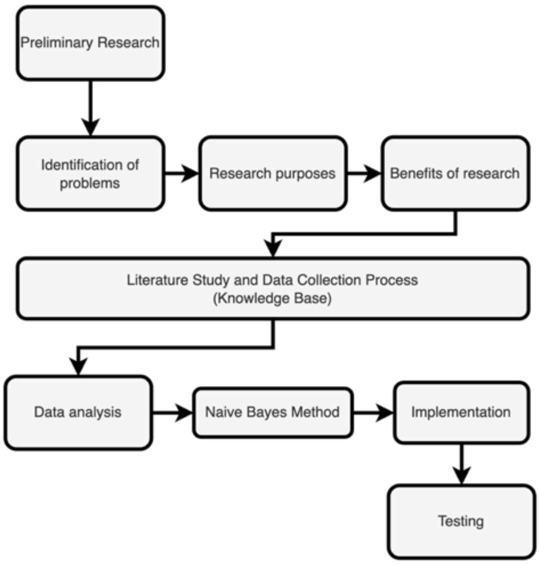

Figure 2. Stages of Research

Figure 2 illustrates the steps in the research and development process of the naïve bayes algorithm. It starts with "Preliminary Research", where the objectives, problems, and benefits of the research are identified. Next is the "Literature Study and Data Collection Process" which serves as a knowledge base. Then, a "Data analysis" is performed, perhaps using statistical methods such as the "Naive Bayes Method". Once the method is implemented, the process ends with "Testing" to verify the effectiveness of the solution that has been developed. It shows the iterative cycle from problem discovery to solution implementation and testing.

Hematochezia refers to the discharge of bright red or dark red blood from the rectum, representing a form of upper gastrointestinal bleeding commonly encountered in the hospital's emergency department. While the majority of patients arrive in a stable condition, some present in urgent situations necessitating prompt and precise interventions [12]. Conditions frequently associated with the symptom of Hematochezia including Hemorrhoids (P01), Inflammatory Bowel Disease (P02), Amoebic Colitis (P03), Colon Malignancy (P04), and Proctitis (P05). Each of these conditions shares the common symptom of Hematochezia, posing a challenge for patients in discerning the specific disease they may be experiencing.

The symptom data for each disease is detailed in table 1.

Table 1. Illness Symptom

Symptom	Name of Symptom
SY01	Pain in the rectal/anal area
SY02	Hematochezia
SY03	Diarrhea
SY04	Constipation
SY05	Inflammation in the rectal/anal area
SY06	Lose weight
SY07	Fatigue
SY08	Cramp or abdominal pain
SY09	Change in bowel habit
SY10	Fever
SY11	Lump in the rectal/anal area
SY12	Itchy and irritated rectal/anal area
SY13	Nausea and vomit
SY14	Hard stool
SY15	Bloated stomach/intestine sensation
SY16	Anal mucus
SY17	Anemia
SY18	Chest and back pain
SY19	Changes in stool color

Following the identification of all disease symptoms, we will proceed to illustrate the correlation between symptoms and diseases, as presented in Table 2.

Table 2. Symptom and Disease Relation

Code	P01	P02	P03	P04	P50
SY01	*	*			*
SY02	*	*	*	*	*
SY03		*	*		*
SY04	*			*	
SY05	*				*
SY06		*		*	
SY07		*		*	
SY08		*	*		*
SY09		*		*	
SY10		*	*		*
SY11	*				
SY12	*				*
SY13		*	*	*	
SY14	*			*	
SY15	*				
SY16		*	*		*
SY17		*		*	
SY18				*	
SY19	*	*	*	*	*

This table classifies and contrasts symptoms among five distinct health conditions: Hemorrhoids (P01), Inflammatory Bowel Disease (P02), Amoebic Colitis (P03), Colon Malignancy (P04), and Proctitis (P50), placing particular emphasis on the symptom hematochezia (SY02). Symptoms common to multiple conditions, like rectal/anal pain (SY01) and hematochezia (SY02), are denoted with an asterisk (*). The table facilitates the identification of both unique and shared symptoms among diseases, such as swelling in the anus area (SY05) linked to

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5064

Hemorrhoids and Proctitis, or diarrhea (SY03) and fever (SY10) more prevalent in Inflammatory Bowel Disease and Amoebic Colitis. Through this approach, your research aims to utilize the Naïve Bayes method for analyzing these symptom patterns, enhancing clinical diagnosis with artificial intelligence, and improving the accuracy of identifying health conditions based on hematochezia symptoms.

Naive Bayes is a straightforward classification method grounded in probability calculations. It explores all potential outcomes by considering diverse combinations and frequency values within existing data. This algorithm applies Bayes' theorem to estimate attributes which assumed to be independent by giving the values of the class variable. Naive Bayes relies on probability and calculations, initially conceptualized by the British scientist Thomas Bayes. The method generates probability predictions based on previous experiences [13].

The procedural steps of the Naive Bayes method are illustrated in Figure 2.

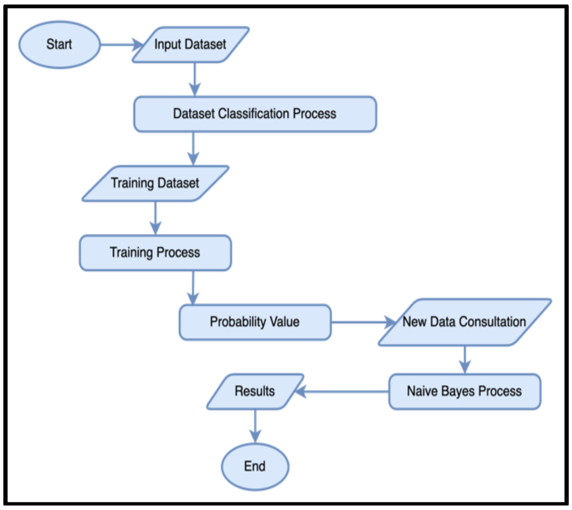

Figure 3. Naïve Bayes Method Flowchart

Flowchart illustrating the Dataset Classification Process for the Expert System Identifying Diseases based on Hematochezia Symptoms using the Naïve Bayes

1. Start: The expert system was initialized by inputting a dataset containing patient data with hematochezia symptoms and other symptoms related to various diseases. This dataset was then classified into two parts: training set and testing set.
2. Input Dataset: The classified dataset was then entered into the expert system.
3. Dataset Classification Process: The training set was used to train the Naïve Bayes model. This model learned to calculate the conditional probabilities of symptoms presented in the training dataset with different disease classes.
4. Training Dataset: The Naïve Bayes model was then tested using the testing set. Using 86% training data and 14% testing data.
5. Probability Value: After the training was complete, the Naïve Bayes model could generate probability values indicating the likelihood of a specific disease based on the exhibited symptoms.
6. New Data Consultation: New data from patients who sought consultation with the expert system was inputted into the system.
7. Naive Bayes Process: In this stage, the Naïve Bayes classification model would learn the relationship between symptoms and the associated diseases. This process was carried out by calculating the probability of symptoms for each disease. The equation used is as follows [14]:

$$P(X|H) = \frac{(P(X|H) * (P(H)}{P(X)} \qquad (1)$$

In this scenario, X represents data with an unspecified class, while H signifies the hypothesis that data X pertains to a particular class. The equation P(H|X) denotes the probability of hypothesis H under the condition X, with P(H) indicating the probability of hypothesis H. Additionally, P(X|H) represents the likelihood of X given the condition of hypothesis H, and P(X) stands for the probability of X.
8. Outcomes: The results of the classification process were showcased, typically manifesting as the most probable disease diagnosis, along with the confidence level or probability determined by the model.
9. Conclusion: The consultation process concluded and the expert system furnished output that healthcare practitioners can leverage for clinical decision-making or to obtain further recommendations.

RESULTS AND DISCUSSION

In this section, we will discuss the implementation of the Expert System for Disease Identification using the Naïve Bayes method for diagnosing diseases with Hematochezia symptoms. The research outcomes will focus on the accuracy, sensitivity, and specificity of the system. The discussion will also address factors influencing the system's effectiveness and compare it with similar systems from other studies. The findings from this research provide crucial insights into the potential and limitations of the system in a broader medical context.

In the process of diagnosis using the Naïve Bayes method, we calculated the probability that the patient suffered from five different diseases based on the symptoms of hematochezia. This probability is called posterior probability.

The posterior probability for each disease is calculated by multiplying the probability of prior of the disease by the probability of the appearance of hematochezia symptoms in the disease. The probability of prior is the probability that the patient suffers from the disease in general, regardless of the visible symptoms. The probability of the appearance of hematochezia symptoms in the disease is the probability that hematochezia symptoms will appear in patients with the disease.

In this case, the total posterior probability for all diseases is 0.00009056. The highest posterior probability was for P01 disease (0.514738), followed by P02 disease (0.292881), P03 disease (0.101413), P04 disease (0.069483), and P05 disease (0.021485). Based on these values, P01 disease is identified as the most likely diagnosis based on the symptoms given. This means that the patient has P01 (Hemorrhoid) disease.

The following are the stages of the Naïve Bayes method in identifying diseases with Hematochezia symptoms.

Prior Probability Stage

Based on the disease and symptom data in Tables 1 and 2, the prior probability values for the existing dataset were obtained and will be presented in Table3.

Table 3. Prior Probability Value

Code	Amount of Case	Probability
P01	50	0.2326
P02	40	0.1860
P03	60	0.2791
P04	30	0.1395
P05	35	0.1628
Total	215	1.0000

Out of the 215 patients included in this study, the disease distribution is as follows: Hemorrhoids (P01) encompass 23.26% of the cases, signifying a noteworthy prevalence. Inflammatory Bowel Disease (P02) was present in 18.60% of patients, indicating its relatively common occurrence. Amoebic Colitis (P3) stands out as the most prevalent disease in this group, representing 27.91% of cases. Conversely, Colon Malignancy (P4) and Proctitis (P5) constitute are 13.95% and 16.28%, respectively, suggesting their comparatively lower occurrence compared to other conditions in this study.

The graph in Figure 4 illustrates the number of patients experiencing symptoms for each disease.

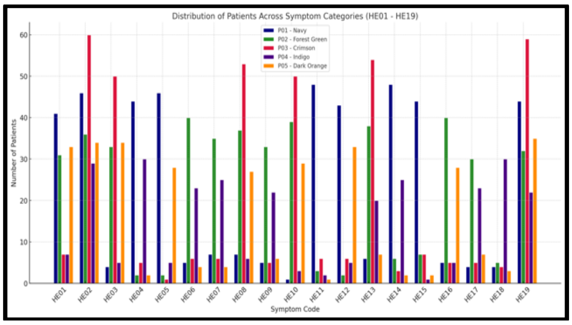

Figure 4. Patient's Illness Symptoms

Figure 4 illustrates that SY01 is most frequently observed in patients with the code P01 (41 cases), followed by P02 (31 cases), P03 (7 cases), P04 (7 cases), and P05 (33 cases). In total, there are 119 instances of SY01. SY02 is also prevalent, with 46 cases in P01, 36 cases in P02, 60 cases in P03, 29 cases in P04, and 34 cases in P05, making a total of 205 cases. SY03, on the other hand, shows a different distribution with 4 cases in P01, 33 cases in P02, 50 cases in P03, 5 cases in P04, and 34 cases in P05, resulting in a total of 126 cases. SY04 is the least commonly observed, with 44 cases in P01, 2 cases in P02, 5 cases in P03, 30 cases in P04, and 2 cases in P05, making a total of 83 cases.

Overall, this data suggests that specific symptoms are more prevalent in patients with particular symptom codes. Subsequently, the probability values for each symptom regarding diseases will be detailed in Table 4.

Table 4. The Value of Disease Symptoms Prior Probabilty

Code	P01	P02	P03	P04	P05
SY01	0,3445	0,2605	0,0588	0,0588	0,2773
SY02	0,2244	0,1756	0,2927	0,1415	0,1659
SY03	0,0317	0,2619	0,3968	0,0397	0,2698
SY04	0,5301	0,0241	0,0602	0,3614	0,0241
SY05	0,5610	0,0244	0,0122	0,0610	0,3415
SY06	0,0641	0,5128	0,0769	0,2949	0,0513
SY07	0,0909	0,4545	0,0779	0,3247	0,0519
SY08	0,0538	0,2846	0,4077	0,0462	0,2077
SY09	0,0704	0,4648	0,0704	0,3099	0,0845
SY10	0,0082	0,3197	0,4098	0,0246	0,2377
SY11	0,8000	0,0500	0,1000	0,0333	0,0167
SY12	0,4831	0,0225	0,0674	0,0562	0,3708
SY13	0,0480	0,3040	0,4320	0,1600	0,0560
SY14	0,5714	0,0714	0,0357	0,2976	0,0238
SY15	0,7213	0,1148	0,1148	0,0164	0,0328
SY16	0,0602	0,4819	0,0602	0,0602	0,3373
SY17	0,0580	0,4348	0,0725	0,3333	0,1014
SY18	0,0870	0,1087	0,0870	0,6522	0,0652
SY19	0,2292	0,1667	0,3073	0,1146	0,1823

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5064

Table 4 displays probability information for diverse symptoms linked to five diseases. These probabilities signify the chances of these symptoms manifesting in the diagnosis scenario of each respective disease. Additional symptoms like pain or discomfort, rectal itching, rectal swelling, etc., manifest with varying probabilities, indicating their prevalence in the diagnostic cases of each disease. These probabilities are derived from the analysis of clinical data and can aid medical practitioners in diagnosing diseases based on the patient's symptom profile.

With these prior probabilities, we can proceed to calculate the posterior probability for the observed symptoms using the Naïve Bayes formula [15]:

$$PP = \frac{\text{Pro Likelihood} \times \text{Pro Prior}}{\text{Total Pro Evidence}} \quad (2)$$

Stage of Likelihood Probability

While Likelihood Probability represents the chance of observing the given symptoms in the presence of the disease, Evidence Probability signifies the overall probability of the observed symptoms across all categories.

After obtaining the prior probability results from the training data, the subsequent step involves conducting the testing process with new user-entered data. This new dataset includes symptoms chosen by the user, specifically [SY01, SY02, SY04, SY07, SY08]. Following this, the Likelihood Probability search process will be executed by multiplying the prior probability values of each symptom for each disease based on the training data. Presented below are the likelihood probability outcomes for each disease:

a. Likelihood for P01:
$0.3445 \times 0.2244 \times 0.5301 \times 0.0909 \times 0.0538 = 0.000200$
b. Likelihood for P02:
$0.2605 \times 0.1756 \times 0.0241 \times 0.4545 \times 0.2846 = 0.000143$
c. Likelihood for P03:
$0.0588 \times 0.2927 \times 0.0602 \times 0.0779 \times 0.4077 = 0.000033$
d. Likelihood for P04:
$0.0588 \times 0.1415 \times 0.3614 \times 0.3247 \times 0.0462 = 0.000045$
e. Likelihood for P05:
$0.2773 \times 0.1659 \times 0.0241 \times 0.0519 \times 0.2077 = 0.000012$

These likelihood values indicate the probability of observing the given symptoms in the presence of a particular disease. However, these values must undergo normalization before they can be understood as posterior probabilities.

Calculating The Posterior Probabilities

The following is a detailed calculation of the posterior probabilities for each disease from P01 to P05:

a. P01:
Likelihood: 0.000200
Prior: 0.2326
Numerator (Likelihood × Prior): 0.000046
Total Evidance =
$0.000200 \times 0.2326 + 0.000143 \times 0.1860 + 0.000033 \times 0.2791 + 0.000045 \times 0.1395 + 0.000012 \times 0.1628 = 0.00009056071361905029$
Posterior: 0.514738
b. P02:
Likelihood: 0.000143
Prior: 0.1860
Numerator (Likelihood × Prior): 0.000027
Total Evidance = 0.000091660713
Posterior: 0.292881
c. P03:
Likelihood: 0.000033
Prior: 0.2791
Numerator (Likelihood × Prior): 0.000009
Total Evidance = 0.000091660713
Posterior: 0.101413
d. P04:
Likelihood: 0.000045
Prior: 0.1395
Numerator (Likelihood × Prior): 0.000006
Total Evidance = 0.000091660713
Posterior: 0.069483
e. P05:
Likelihood: 0.000012
Prior: 0.1628
Numerator (Likelihood × Prior): 0.000002
Total Evidance = 0.000091660713
Posterior: 0.021485

Following the completion of the Bayesian calculation, the subsequent step involves translating the expert system into a web-based program. The results of this implementation are shown in Figure 5.

Figure 5. Initial display of the program

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v9i2.5064

Figure 5 displays the main page of the expert system, emphasizing its consultation service for symptom-based Hematochezia diagnosis. It highlights two features: selecting immediate consultation and accessing consultation history. This illustrates the system's capability for direct user interaction and monitoring consultation records.

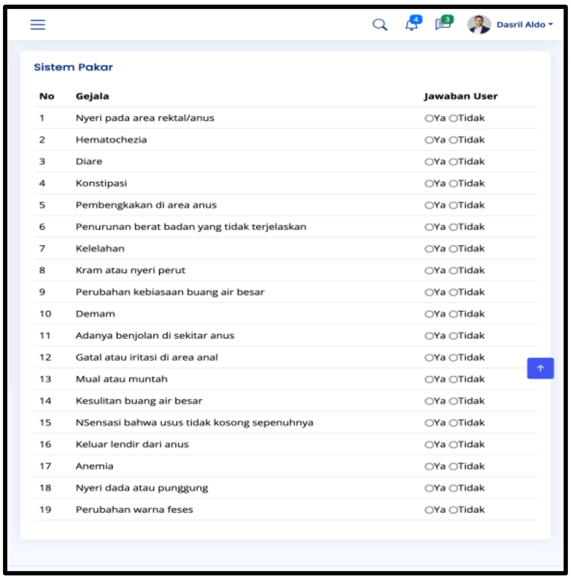

Figure 6. Consultation form display.

Figure 7 presents the consultation form where one can choose the symptoms observed by the patient, and these symptoms are selec Table from the provided list.

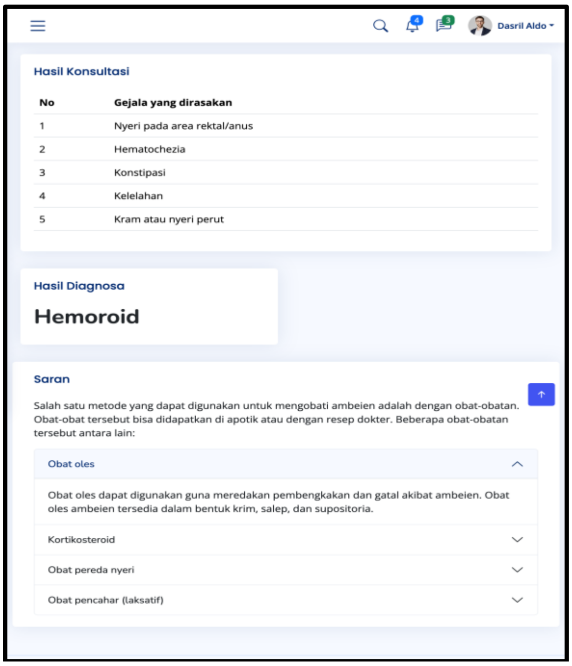

Figure 7. Display of diagnostic results

Upon entering the symptoms, the software will present the diagnostic findings, revealing the probability of the patient having a particular illness, along with the associated confidence or probability level. The expert system will then offer treatment suggestions based on the diagnostic results, encompassing potential interventions such as medication, medical procedures, or adjustments to one's lifestyle.

From Figures 5, 6 and 7 are the results of the implementation of the expert system, where Figure 5 shows the start page of the expert system consisting of the title and main menu of the expert system program. Then Figure 6 shows a consultation page containing symptoms of Hematochezia Symptom-Based Disease that can be chosen by patients according to the conditions experienced. Furthermore, Figure 7 shows the results of consultation from patients containing symptoms that have been selected, diagnosis results, suggestions that can be done by patients.

The next phase involves examining the outcomes of the expert system in comparison to the diagnoses given by professionals, as illustrated in Table5.

Table 5. Test Result

Code	Expert System	Expert	Result
Case 1	Inflammatory Bowel Disease	Inflammatory Bowel Disease	✓
Case 2	Proctitis	Proctitis	✓
Case 3	Proctitis	Proctitis	✓
Case 4	Proctitis	Proctitis	✓
Case 5	Hemoroid	Amoebic Colitis	✗
Case 6	Proctitis	Proctitis	✓
Case 7	Colon Malignancy	Colon Malignancy	✓
Case 8	Proctitis	Proctitis	✓
Case 9	Colon Malignancy	Colon Malignancy	✓
Case 10	Inflammatory Bowel Disease	Inflammatory Bowel Disease	✓
Case 11	Inflammatory Bowel Disease	Inflammatory Bowel Disease	✓
Case 12	Colon Malignancy	Colon Malignancy	✓
Case 13	Inflammatory Bowel Disease	Inflammatory Bowel Disease	✓
Case 14	Colon Malignancy	Colon Malignancy	✓
Case 15	Colon Malignancy	Colon Malignancy	✓
Case 16	Amoebic Colitis	Amoebic Colitis	✓
Case 17	Amoebic Colitis	Amoebic Colitis	✓
Case 18	Amoebic Colitis	Amoebic Colitis	✓
Case 19	Amoebic Colitis	Amoebic Colitis	✓
Case 20	Inflammatory Bowel Disease	Inflammatory Bowel Disease	✓

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5064

Code	Expert System	Expert	Result
Case 21	Hemoroid	Hemoroid	✓
Case 22	Proctitis	Proctitis	✓
Case 23	Proctitis	Proctitis	✓
Case 24	Amoebic Colitis	Amoebic Colitis	✓
Case 25	Proctitis	Proctitis	✓
Case 26	Hemoroid	Hemoroid	✓
Case 27	Hemoroid	Hemoroid	✓
Case 28	Hemoroid	Hemoroid	✓
Case 29	Proctitis	Proctitis	✓
Case 30	Proctitis	Proctitis	✓
Case 31	Hemoroid	Colon Malignancy	X
Case 32	Amoebic Colitis	Amoebic Colitis	✓
Case 33	Hemoroid	Hemoroid	✓
Case 34	Colon Malignancy	Colon Malignancy	✓
Case 35	Hemoroid	Hemoroid	✓

Analyzing the test outcomes, there were 2 (two) discrepancies. The expert system identified Hemorrhoids while the expert diagnosed Amoebic Colitis in Case 5. Meanwhile, for Case 31, the expert system identified Hemorrhoids while the expert diagnosed Colon Malignancy.

To find the accuracy value of the expert system using the Confusion Matrix which can be seen in table 6.

Table 6. Confusion Matrix

	Expert System +	Expert System -
Expert +	32	0
Expert -	2	1

From table 6 there are 32 data that get the same diagnosis between doctors and expert systems. It is located in the section "Expert System (+) / Doctor (+)". There are 2 different diagnostic data between doctors and expert systems. It is located in the section "Expert System (-) / Doctor (+)". There is 1 data that gets the correct diagnosis by the expert system but not by the doctor. It is located in the section "Expert System (+) / Doctor (-)". The "Expert System (-) / Doctor (-)" section states that one data is not correctly diagnosed by both systems. Next calculate these values, where Accuracy = 35/33 = 0.9429, Precision = 32/32 = 1, Recall = 32/34 = 0.9412, F-1 Score = 2×1×0.9412/1+0.9412 = 0.9706.

Based on the results of the expert system performance evaluation, it can be concluded that the system shows excellent performance in making accurate predictions. With an accuracy value of about 94.29%, a precision of 100%, and a recall of about 94.12%, the expert system is able to provide a diagnosis consistent with the doctor in most cases. A high F-1 score, approximately 97.06%, indicates a balanced precision and recall in the system.

CONCLUSION

The findings of this study suggest that the expert system, which was created using the Naïve Bayes method, proves to be effective in assisting the diagnosis of diseases characterized by Hematochezia symptoms with a noTablelevel of accuracy. The system has demonstrated its proficiency in recognizing ailments like Hemorrhoids, Inflammatory Bowel Disease, Amoebic Colitis, Colon Malignancy, and Proctitis. This examination is carried out by considering symptoms chosen by the user, including discomfort in the rectal region, Hematochezia, constipation, fatigue, and abdominal cramps or pain. The outcome is a diagnosis of Hemorrhoids with a posterior probability value of 0.514738.
During a test with 35 case samples, two disparities were observed between the diagnostic outcomes of the expert system and those of the expert. Nevertheless, the system exhibited an impressive level of accuracy, approximately 94.29%. The expert system also shows significant potential as a reliable tool for diagnosing diseases based on Hematochezia symptoms, provide important support for patients and health practitioners in making the right clinical decisions.

REFERENCE

[1] J. Zhao, Q. Li, Y. Gong, and K. Zhang, “Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks,” *IEEE Trans Veh Technol*, vol. 68, no. 8, pp. 7944–7956, Aug. 2019, doi: 10.1109/TVT.2019.2917890.

[2] H. Mao, M. Schwarzkopf, S. B. Venkatakrishnan, Z. Meng, and M. Alizadeh, “Learning scheduling algorithms for data processing clusters,” in *Proceedings of the ACM Special Interest Group on Data Communication*, New York, NY, USA: ACM, Aug. 2019, pp. 270–288. doi: 10.1145/3341302.3342080.

[3] D. S. W. Ting *et al.*, “Artificial intelligence and deep learning in ophthalmology,” *British Journal of Ophthalmology*, vol. 103, no. 2, pp. 167–175, Feb. 2019, doi: 10.1136/bjophthalmol-2018-313173.

[4] T. S. Rachmawati and M. Agustine, “Keterampilan literasi informasi sebagai upaya pencegahan hoaks mengenai

informasi kesehatan di media sosial," *Jurnal Kajian Informasi & Perpustakaan*, vol. 9, no. 1, p. 99, Jun. 2021, doi: 10.24198/jkip.v9i1.28650.

[5] K. Ain, H. B. Hidayati, and O. Aulia Nastiti, "Expert System for Stroke Classification Using Naive Bayes Classifier and Certainty Factor as Diagnosis Supporting Device," *J Phys Conf Ser*, vol. 1445, no. 1, p. 012026, Jan. 2020, doi: 10.1088/1742-6596/1445/1/012026.

[6] G. Irfansyah, U. Darusallam, and B. Benrahman, "Early Diagnosis Expert System Hepatitis Using Naive Bayes Method: Early Diagnosis Expert System Hepatitis Using Naive Bayes Method", Mantik, vol. 3, no. 4, pp. 182-187, Feb. 2020.

[7] B. Budianto, I. Fitri, and W. Winarsih, "Expert System for Early Detection of Disease in Corn Plant Using Naive Bayes Method: Expert System for Early Detection of Disease in Corn Plant Using Naive Bayes Method," J. Mantik, vol. 3, no. 4, pp. 308–317, 2020.

[8] I. Santiko and I. Honggo, "Naive Bayes Algorithm Using Selection of Correlation Based Featured Selections Features for Chronic Diagnosis Disease," *IJIIS: International Journal of Informatics and Information Systems*, vol. 2, no. 2, pp. 56–60, Sep. 2019, doi: 10.47738/ijiis.v2i2.14.

[9] S. K. Maliha, R. R. Ema, S. K. Ghosh, H. Ahmed, Md. R. J. Mollick, and T. Islam, "Cancer Disease Prediction Using Naive Bayes,K-Nearest Neighbor and J48 algorithm," *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–7, Jul. 2019, doi: 10.1109/ICCCNT45670.2019.8944686.

[10] T. Siahaan, Y. Laia, M. Silitonga, C. Pasaribu, F. Sains, and D. Teknologi, "Penerapan Data Mining Classification Untuk Data Pasien Covid-19 Menggunakan Metode Naïve Bayes," *AJurnal TEKINKOM*, vol. 6, no. 1, 2023, doi: 10.37600/tekinkom.v6i1.879.

[11] M. T. Hayat Suhendar and Y. Widyani, "Machine Learning Application Development Guidelines Using CRISP-DM and Scrum Concept," in *2023 IEEE International Conference on Data and Software Engineering (ICoDSE)*, IEEE, Sep. 2023, pp. 168–173. doi: 10.1109/ICoDSE59534.2023.10291438.

[12] Darmadi Darmadi and Sania Audry Nasution, "Perdarahan Saluran Cerna Atas", Termometer, vol. 2, no. 1, pp. 193–207, Nov. 2023.

[13] R. Rachman, R. N. Handayani, and I. Artikel, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," J. Inf., vol. 8, no. 2, pp. 111–122, 2021.

[14] R. Rachman, "Sistem Pakar Deteksi Penyakit Refraksi Mata Dengan Metode Teorema Bayes Berbasis Web," J. Inform., vol. 7, no. 1, pp. 68–76, 2020.

[15] S. Lestari, Akmaludin, and M. Badrul, "Implementasi Klasifikasi Naive Bayes Untuk Prediksi Kelayakan Pemberian Pinjaman Pada Koperasi Anugerah Bintang Cemerl," *PROSISKO*, vol. 7, no. 1, pp. 8–16, Mar. 2020.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri