Indonesian Journal of Electrical Engineering and Computer Science
Vol. 40, No. 2, November 2025, pp. 829~839
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v40.i2.pp829-839

Application of naïve bayes algorithm in expert system for diagnosing chilli plant diseases based on growth phase on peatland

Fatayat, Wahyu Lestari, Alfirman
Department of Computer Science, Universitas Riau, Pekanbaru, Indonesia

Article Info

Article history:

Received Sep 17, 2024
Revised Aug 5, 2025
Accepted Oct 14, 2025

Keywords:

Chilli plants
Expert system for
Peatland disease diagnosis
Naive Bayes

ABSTRACT

Agricultural development on peatlands has its own challenges, especially in the cultivation of chili plants that are susceptible to various diseases. Therefore, an expert system is needed that can help farmers diagnose chili plant diseases quickly and accurately based on the plant growth phase. This research aims to apply the Naïve Bayes algorithm to the expert system for diagnosing Capsicum annum L (Chilli) plant diseases. The results of the expert system research offer an innovative and adaptive solution for the management of plant diseases in peatlands, with great potential to increase agricultural productivity and plant resistance to disease. The expert system is able to diagnose several types of diseases on chili plants in peatlands, such as anthracnose, fusarium wilt, and leaf curl disease. Each diagnosis is based on symptoms observed in each phase of plant growth, from the vegetative phase to the generative phase. Expert system testing results. This system is expected to increase the productivity and quality of chili crops on peatlands, as well as reduce losses due to disease attacks. In addition, this research also shows that the Naive Bayes algorithm has great potential to be applied in expert systems in other agricultural fields.

Corresponding Author:

Fatayat
Department of Computer Science, Universitas Riau
Pekanbaru, Indonesia
Email: fatayat@lecturer.unri.ac.id

1. INTRODUCTION

The chili plant (Capsicum annanum) is a plant that has the characteristics of red fruit when it is ripe and green when it is unripe and has a very spicy taste. This spicy taste is caused by the capsaicin substance contained in the chili fruit. Chili plants in Indonesia are usually cultivated in former rice fields or dry plantation areas. Plant growth requirements must be fulfilled so as to obtain a lot of fruit and plants that grow lushly. Chilies usually grow in various types of soil that have irrigation and sufficient water availability during the growth and development of the chili plant [1]. The development of information technology (IT) is now very rapid and has been used in various aspects of life, both in the fields of government, banking, socio-culture, industry, education, and even agriculture. This IT development is very helpful for human work, both in terms of ease of work, efficiency, work completion time, and accuracy of work results. There are still many farmers who do not know how to properly care for chili plants, so that the harvest is less than what should be obtained. To produce good chili plants, farmers must pay attention to care, starting with seed care, good land, and suitable fertilizers according to plant conditions. The application of the Naive Bayes algorithm in the expert system for diagnosing chili plant diseases based on the web-based growth phase is a significant step in increasing the efficiency and accuracy of diagnosis to evaluate the symptoms of chili plant diseases. One of the factors that causes the low quality of chili is that it is threatened by pests and plant diseases. Pest and disease attacks in the growth phase cause chili plants to die before they grow and cannot be harvested. The Naïve Bayes method is used in diagnosing diseases in chili plants, where each alternative provided will be ranked to get the best results. Therefore, the author wants to design a web-based system. According to previous research from [2].

Expert system to diagnose diseases in corn plants with the Bayes Method, where the method used in diagnosing diseases in corn plants is the Bayes Method, where each alternative provided will be ranked to get the best results [3]. The results produced are in the form of ranking data on diseases in corn plants, which are used as a tool in decision-making for farmers. This algorithm was chosen because of its ability to handle data with attributes that are independent of each other, as well as its simplicity in implementation and interpretation of results [4], [5]. Irsyada *et al.* [6]. This approach allows for a more accurate diagnosis, as the symptoms and type of disease affecting a plant can vary depending on the growth phase of the plant.

The main innovation is the use of a plant growth phase-based approach for disease diagnosis. This model takes into account that the symptoms and diseases affecting chili plants can be different in each growth phase (e.g., vegetative, generative, to fruit formation phase). This provides a more accurate diagnosis because each growth phase has a different risk of disease [4]. In this study, a specialized expert system was optimized for the unique conditions of peatlands, which have characteristics such as high soil acidity, high humidity, and limited soil aeration. This innovation enables faster and more accurate diagnosis of plant diseases and is adaptive to changing environmental conditions in real-time [7], [8]. Integrated sensor data provides a more comprehensive picture of plant health. One of them is the development of web-based or mobile applications that are intuitive for users (farmers) to diagnose diseases directly based on the input of observed symptoms. This innovation makes it easier for farmers to utilize advanced diagnostic technology without the need for in-depth technical knowledge. According to previous research on 'implementation of the Naive Bayes classifier algorithm as a thesis supervisor recommendation system'. In this study, 217 reference training data for supervisor 1 and 177 training data for supervisor 2 were used [9]. While the test data used was 10 data. The criteria used are competence, functional position, and the lecturer's home base. The implementation of the Naive Bayes Classifier algorithm is inserted in the SIMASITA CIC integrated thesis submission system application. Based on the test results of the Naive Bayes Classifier algorithm, the early prediction and evaluation of disease severity are extremely important for patient prognosis. But in this paper, we propose to investigate the Naïve Bayes event model for an expert system to diagnose COVID-19 because it has not been considered for this problem before. In response to the outbreak, we summarize the current knowledge of COVID-19 and compare it with previous experiences of the SARS outbreak in Hong Kong, studying effective measures to control the COVID-19 epidemic [10]. While the test data used was 10 data.

The criteria used are competence, functional position, and lecturer's home base. The implementation of the Naive Bayes Classifier algorithm is inserted in the SIMASITA CIC integrated thesis submission system application. Based on the test results of the Naive Bayes Classifier algorithm, the comparison of the suitability level of supervisor 1 is 90%: 10%, and the comparison of the suitability level of supervisor 2 is 30%: 70%. According to previous research from [11]. explains that although the Naïve Bayes algorithm assumes feature independence, it can still achieve high accuracy even when this assumption is violated. The study shows that Naïve Bayes remains effective because it only needs to produce the correct decision boundary rather than perfectly estimating probabilities. As a result, Naïve Bayes is still optimal and performs well in many real-world applications, especially when working with small datasets and high-dimensional features [12]. developed an academic information system that uses the Naïve Bayes algorithm to automatically determine thesis supervisors for students. The system evaluates lecturer suitability based on criteria such as research expertise, academic competence, and workload. Their results show that Naïve Bayes is effective in recommending supervisors who match student research topics and academic needs. The study concludes that this method improves the accuracy and efficiency of supervisor selection in universities, compared to manual assignment [13]. Naïve Bayes Classification for Action Promotion Strategy Recommendations at Aesthetic Dental Clinic Karawang. The Naïve Bayes algorithm can be used to predict customer interest based on promotion, provided by the ADC Clinic (Aesthetic Dental Clinic). One data mining classification algorithm that can be used to support effective and efficient promotional strategies is Naïve Bayes. The results of this study show that application of the Naïve Bayes algorithm can provide important information, such as prediction results, in an effort to attract customer interest [14]. Developed an expert system for pepper disease diagnosis combining Certainty Factor and Naïve Bayes. The system calculates disease probability based on symptoms and provides treatment recommendations. Demonstrates Naïve Bayes effectiveness in horticultural disease diagnosis [15]. Proposes an optimized hybrid approach between Naïve Bayes and rules-based logic to diagnose chili plant diseases. Shows that combining probabilistic and expert-knowledge models improves diagnostic accuracy [16]. Comprehensive review of AI methods for crop disease detection, including Naïve Bayes. Highlights Naïve Bayes as a suitable classifier for early-stage agricultural diagnosis systems with limited training data [17]. Uses AI-based classification to identify pepper leaf diseases. Although deep learning is used, the study supports the importance of dataset construction and symptom identification applicable to Naïve Bayes-based expert systems. Examines early disease detection in chili plants with hyperspectral data. Identifies physiological indicators at different growth phases, serving as relevant input for Naïve Bayes-based plant disease modeling [8]. Improves Naïve Bayes classification accuracy for plant diseases using fuzzy discretization. Demonstrates how handling uncertainty in agricultural symptom data increases performance—relevant for peatland variability [18]. Implements a Bayesian-based expert system to diagnose crop diseases. Emphasizes symptom weighting, probability calculation, and real-time diagnostic support—aligned with chili diagnosis system design [19]. This paper uses the Naïve Bayes algorithm to classify diseases in chili plants based on leaf image features such as texture and color. The Naïve Bayes model achieved good classification performance and proved effective for detecting common chili diseases, demonstrating its suitability for agricultural expert systems [20]. This study develops an expert system using a Naïve Bayes classifier to identify diseases in chili plants. The algorithm performed reliably using symptom-based inputs, confirming Naïve Bayes as a good solution for knowledge-based agricultural diagnostic systems, especially for farmers with limited technical knowledge [21]. Rish conducted experimental analysis on the Naïve Bayes algorithm using multiple datasets. The results show that Naïve Bayes performs efficiently with small training data and works well for various classification problems, particularly text and medical diagnosis, but its performance decreases with highly correlated attributes [22]. This paper evaluates Naïve Bayes in text classification tasks such as spam filtering and sentiment analysis. The results demonstrate the algorithm's strength in handling high-dimensional word features, delivering high accuracy with low computational cost, making it suitable for NLP applications [23]. This study develops a classification model for diseases and pests affecting corn plants (via digital images) by integrating a fuzzy-discretization approach with the Multinomial Naïve Bayes (MNB) classifier. Continuous predictor variables (RGB image channels) are discretised using fuzzy membership functions, allowing overlapping intervals to better capture vagueness in the data. The authors hypothesize that using different combinations of fuzzy membership functions and class interval counts will affect model performance. They conduct experiments with Monte Carlo resampling to evaluate generalisation and compare crisp discretization vs fuzzy discretization within the MNB framework [24].

2. METHOD

The methodology used in this research is 'an expert system for diagnosing Capsicum annum L. (Chili) plant diseases based on growth phase using Naive Bayes algorithm'. research method in collecting data on chili plant diseases based on growth, there are 9 data on Capsicum annum L. Plant diseases based on the growth phase, which can be seen in Table 1.

Table 1. Capsicum annum l. plant disease based on growth phase

Kode	Disease name
P101	Bacterial wilt, Sprout lodging or damping off
P102	Root swollen nematode, Yellow striped mosaic or chlorosis
P103	Bacterial spot
P104	Anthracnose, Cercospora leaf spot
P105	Phytophthora leaf rot
P106	Fusarium wilt
P107	Choanephora leaf rot
P108	Staphylium grey spot
P109	Dwarfing, necrosis and light mosaic

2.1. Data collection stage

Based on current activities at the Department of Agriculture and Food Crops of Rokan Hilir Regency, Riau, at the Office of the Agricultural Extension Office (BPP) of Tanah Putih District, Sedinginan, where to diagnose diseases in chili plants, farmers do several things, such as collecting chili plants that are affected by disease, identifying, then making reports to officers to be able to provide drugs used to overcome chili plant disease problems.

The author applies several research methods in data collection as:

- Literature study: where literature study is needed to obtain additional literature from books, journals, and articles related to the field of expert system science and related to the science of chili plants.
- Observation is the second step in collecting data after the author conducts a literature study.
- Interview: interviews are the next step after observations are made. An interview is a data collection technique by meeting face-to-face directly between the interviewer and the informant. Observation activities and interviews will be carried out at the department of agriculture and food crops of Rokan Hilir regency, Riau.

2.2. Research steps

The research steps in this study were systematically designed to ensure the validity and reliability of the results. The stages begin with data collection and pre-processing, followed by model development, training, and evaluation. The entire process aims to produce accurate findings that align with the research objectives while also contributing to the development of science in related fields. The preparatory steps taken in completing this research can be seen in Figure 1.

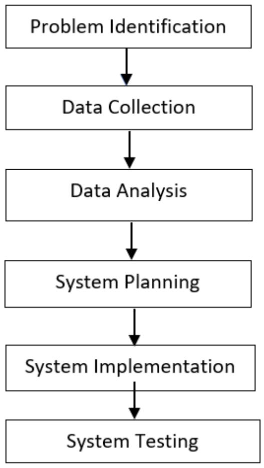

Figure 1. Research steps

2.3. Naive Bayes classifier

Naive Bayes classifier is a simple probability classifier based on Bayes' Theorem. Bayes' Theorem is combined with “naive,” which means that each attribute/variable is independent. The Naive Bayes classifier [4]. Naive Bayes classifier (NBC) algorithm Naive Bayes is one of the machine learning methods that uses probability calculations. This algorithm utilizes the probability and statistical method proposed by british scientist Thomas Bayes, which predicts future probabilities based on previous experience.

2.4. Naïve Bayes algorithm

How Naive Bayes works The Naive Bayes classifier (NBC) algorithm works the workings of the Naïve Bayes classifier are through two stages, namely: learning Naïve Bayes is a supervised learning algorithm, so it requires prior knowledge to make decisions.

2.5. System analysis

In the analysis process using the naive Bayes algorithm, the author uses expert rules obtained from interviews with the department of agriculture and food crops of Rokan Hilir regency, Riau. The expert rules that will be applied to the system that will be built are nine expert rules consisting of conditions and diagnoses.

2.6. System design

System design for making a disease diagnosis system for capsicum annum L. (Chili) plants using visual paradigm CE. In UML design, there are several diagrams needed in the process of making it, namely use case diagrams, activity diagrams, sequence diagrams, and class diagrams. In the process, naive 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.

Using Bayes' theorem, in (1) can be written as

$$n + c + m.p$$
$$P(ai|vj) = n + m \tag{1}$$

Description:
P(ai|vj) = probability of the selected attribute if the disease type is known nc = number of records in the learning data

p = 1 / number of class disease types
m = number of parameter ymptoms
n = number of records in the learning data / each class
Using Bayes' theorem, in (1) can be written as

$$n + c + m.p$$
$$P(ai|vj) = n + m \quad (2)$$

$$VMAP = P(Vj\ | a1, a2, . a3, . an) \quad (3)$$

Description:
VMAP = highest probability
a1, a2, a3, ...an = input attributes

3. RESULT AND DISCUSSION

3.1. System analysis

In the analysis process using the naive Bayes algorithm, the author uses expert rules obtained from interviews with the department of agriculture and food crops of Rokan Hilir regency, Riau, at the agricultural extension office (BPP) tanah putih district, sedinginan. The following expert rules that will be applied to the system that will be built can be seen in Table 2.

In Table 1 above, trials were conducted by obtaining data based on the form of symptoms of plant disease experienced by several farmers. The trial data used 23 symptoms of disease in chili plants using the naive Bayes algorithm. The author uses expert rules obtained from interviews with the Agriculture Office based on the symptoms of chili plant disease. We will mark with a check the system that will be used to prove the calculation process carried out manually, and calculate the system built. The goal is to prove the calculation using the Naive Bayes classifier algorithm. Then the calculation using the Naïve Bayes classifier algorithm can be applied to the disease symptom trial in Table 3. In Table 3, it can be seen the results of symptom codes and names of disease symptoms in chili plants. There are five symptom codes and five disease symptoms found in codes F102, G104, G112, G114, and G122.

Table 2. Expert rules

Condition	Diagnosis								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
G101	√	-	-	-	-	√	-	-	-
G102	-	√	-	-	-	√	-	-	-
G103	√	-	-	-	-	-	-	-	√
G104	-	√	-	√	-	-	-	-	-
G105	-	-	-	-	-	-	-	√	-
G106	-	-	√	-	-	-	-	√	-
G107	-	-	-	√	-	-	-	-	√
G108	-	-	√	√	-	-	-	√	-
G109	-	-	√	-	-	-	-	√	-
G110	-	√	-	-	√	-	√	-	-

Source: Expert rules obtained from interviews with the Department of Agriculture and Food Crops of Rokan Hilir Regency, at the Agricultural Extension Office (BPP) Tanah Putih District, Sedinginan (2023).

Table 3. Disease symptom trial data

Symptom code	Name of disease symptom
G102	Wilting is visible on the lower leaves. After a few days, the entire leaf becomes permanently wilted, while the leaf color remains green, sometimes slightly yellowish. The vascular tissue of the lower stem and roots becomes brownish. If the stem or roots are cut crosswise and dipped in clear water, a cloudy liquid of bacterial colonies will come out, floating in the water like a puff of smoke.
G104	Seedlings are slightly yellowish but often look like healthy plants. There are root nodules that cannot come off even if the roots are rubbed harder. Chlorotic or yellow mottled leaf color.
G112	In the field, plants wilt sporadically. Yellowish leaves and wilting that starts from the upper leaves The wilting occurs gradually until permanent wilting occurs some time later and the leaves remain attached to the stem.
G114	The rot spreads to the lower part of the plant and re-invades the new growth points so that almost all the shoots drip. The stems affected by this disease become dry, rotten, and peel off easily. At high humidity, black hairs form on infected tissues.
G122	Growth is stunted. The distance between petioles is shorter, especially at the top, so that the leaves accumulate and clump together like crackers. Leaves fall so that only a chain of leaves remains, with leaves curled up at the tip of the shoot.

3.2. Naive Bayes classifier calculation
3.2.1. Determine the NC value for each class

The determination of the NC value is based on the selection of symptoms found in a disease. The first number 1 states that the symptom is present in the disease, while the number 0 does not.

1st disease: bacterial wilt, sprout lodging, or damping off n = 1.
p = 1/9 = 0,111
m = 23 G102. $n_c = 0$
G104. $n_c = 0$
G112. $n_c = 0$
G114. nc = 0
G122. nc = 0
And disease: root swollen nematode, yellow striped mosaic, or chlorosis n = 1
p = 1/9 = 0,111 m = 23
G102. nc = 1
G104. nc = 1
G112. nc = 0
G114. nc = 0
G122. nc = 0
3rd disease: bacterial spot n = 1
p = 1/9 = 0,111 m = 23
G102. nc = 0
G104. nc = 0
G112. nc = 0
G114. nc = 0
G122. nc = 0

determine the classification result, namely vi, which has the largest multiplication result. The result of vi, which has the largest multiplication using in (3), is Obtained in Table 4.

Table 4. Comparison of v value of sample classification results

Code	Disease	Nilai
P101	Bacterial wilt, sprout lodging, or damping off	1,52
P102	Root, swollen nematode, yellow, striped mosaic, or chlorosis	2,94
P103	Bacterial spot	1,52
P104	Anthracnose, Cercospora leaf spot	2,94
P105	Phytophthora leaf rot	2,94
P106	Fusarium wilt	2,94
P107	Choanephora leaf rot	2,11
P108	Staphylium gray spot	1,52

1st disease: bacterial wilt, sprouting lodging, or damping off

$P(G102|P1)=\frac{0+23\text{ x }0{,}111}{1+23}$

$P(G102|P1)=0{,}10648$

$P(G104|P1)=\frac{0+23\text{ x }0{,}111}{1+23}$

$P(G102|P1)=0{,}10648$

$P(G112|P1)=\frac{0+23\text{ x }0{,}111}{1+23}$

$P(G102|P1)=0{,}10648$

$P(G114|P1)=\frac{0+23\text{ x }0{,}111}{1+23}$

$P(G102|P1)=0{,}10648$

$P(G122|P1)=\frac{0+23\text{ x }0{,}111}{1+23}$

$P(G102|P1)=0{,}10648$

Calculate P(vj) × P(ai|vj) for each vi using equation 2.2 1st disease: Bacterial wilt, sprout lodging, or damping off

P(P1) x [P(G102|P1) x P(G104|P1) x P(G112|P1) x P(G114|P1) x P(G122|P1)]
= 0.111 x 0,10648 x 0,10648 x 0,10648 x 0,10648 x 0,10648
= 1,521E-06

2nd disease: root swelling nematode, yellow mottled mosaic, or chlorosis

P(P2) x [P(G102|P2) x P(G104|P2) x P(G112|P2) x P(G114|P2) x P(G122|P2)]
= 0.111 x 0,14815 x 0,14815 x 0,10648 x 0,10648 x 0,10648
= 2,94423E-06

Since the value of 2.94 is the highest and there are 4 diseases with the same value, the case example for plant diseases in farmer 1 (one) is classified as disease type ‘P106 - Phytophthora late blight’, and the other 3 recommendations are P105 - Phytophthora late blight, P104 - Anthracnose, Cercospora leaf spot, P102 - Root swollen nematode, Yellow striped mosaic, or chlorosis. P106 was chosen first because it has the highest disease code of the other 3 recommendations.

3.3. Use case diagram

The use case diagram design will display how someone (actor) utilizes the system or uses the existing system. In the system to be built, there is 1 main actor, namely the admin/operator, and 1 supporting actor, namely the user or farmer/visitor. A use case diagram can be seen in Figure 2.

In the use case diagram above, the use case diagram design will display how a person (actor) utilizes the system or uses the existing system to be built. There are two actors, namely 1 main actor, namely the admin or operator, who can run the system in the login section and enter the main menu to do, input data, edit data, and save data, and the second actor as a visitor can see disease symptoms and print disease diagnosis results.

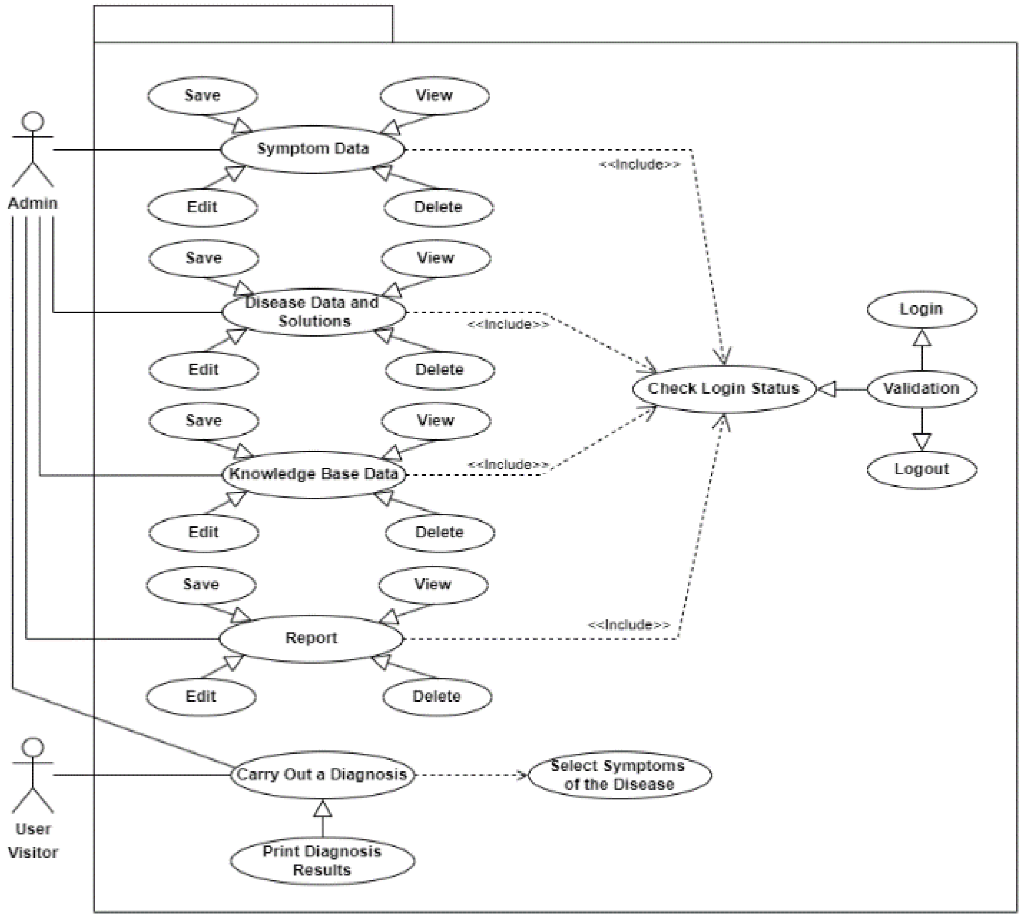

Figure 2. Use case diagram

3.4. System implementation

At this stage is the process of making a system that refers to the results of system analysis and design. Calculations using the Naïve Bayes algorithm will be implemented in the system.

3.5. Login page display

The login page view can be seen in Figure 3. In the picture, the login page is a page where users can input their username and password to be verified and given access rights to the system. The login page is intended for admin as a system user.

Figure 3. Login page display

3.6. Display of diagnosis results page

After the user selects the symptoms, the system will display the diagnosis results containing the user's name, disease name, and solution or control. The display of the diagnosis results page can be seen in Figure 4. From the diagnostic results of Figure 4 above, the classification is vi, which has the largest multiplication result. The result of vi, which has the largest multiplication using equation 2.3, obtained a value of 2.944. Because the value of 2.94 is the largest and there are 4 diseases with the same value, then the case example for plant disease in the 1st farmer (one) is classified as the type of disease ‘P106 - Phytophthora leaf rot’, and the other 3 recommendations are P105 - Phytophthora leaf rot, P104 - Anthracnos, Cercospora leaf spot, P102 - Root swollen nematode, Yellow striped mosaic, or chlorosis. P106 was chosen as the most important because it has a disease code.

Figure 4. Display of diagnosis results page

4. CONCLUSION

The application of an expert system to diagnose plant diseases of capsicum annum L. (Chili) plants based on the growth phase using the naive Bayes algorithm can be one solution to help the department of agriculture and food crops of Rokan Hilir regency, Riau. The Naïve Bayes algorithm has proven effective in identifying and diagnosing chili plant diseases in various growth phases and is able to achieve an adequate level of accuracy in detecting diseases in chili plants. The results of the research data on the symptoms of chili plant disease on peatlands in the calculation of Naive Bayes determine the classification result is vi, which has the largest multiplication result. Comparison of the value v of the sample classification results obtained a value of 2.94 is the largest and there are 4 diseases with the same value, so the case example for plant disease in the 1st farmer (one) is classified as the type of disease 'P106 **-** Phytophthora leaf rot', and the other 3 recommendations are P105 **-** Phytophthora leaf rot, P104 **-** Anthracnose, Cercospora leaf spot, P102 **-** Root swollen nematode, Yellow striped mosaic, or chlorosis. P106 was chosen foremost because it has a disease code.

Although this research focuses on chili plant diseases on peatlands and is web-based, it can be further developed using other crops grown on peatlands or other soil types, and the system can be developed into Android and mobile applications. Similar applications can be developed for various agricultural commodities, helping farmers diagnose diseases in various types of plants. Peatlands can be developed due to their characteristics, such as high moisture levels, low soil pH, and tendency towards fire or rapid chemical changes. This research also has prospects for collaboration with agricultural research institutions and the government in order to develop technologies to increase agricultural productivity. using this system as part of a smart agriculture support program for farmers in various regions.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to all those who have provided personal assistance in the completion of this work. Special thanks are extended to the technical advisors who provided valuable input during the design of the methodology, and to the language editors who helped improve the manuscript for clarity and readability. All parties mentioned in this section have been informed and have given their consent to be named.

FUNDING INFORMATION

This research was financially supported by the Budget Implementation List (DIPA) of the University of Riau. The authors gratefully acknowledge this financial support, which enabled the completion of the work presented in this article.

AUTHOR CONTRIBUTIONS STATEMENT

This journal adopts the Contributor Roles Taxonomy (CRediT) to clearly recognize and document individual author contributions. In this work, **Fatayat** contributed to conceptualization, methodology, software, validation, formal analysis, investigation, writing original draft, writing review and editing, visualization, and funding acquisition. **Wahyu Lestari** contributed to methodology, investigation, writing original draft, writing review and editing, visualization, supervision, and project administration. **Alfirman** contributed to conceptualization, software, validation, data curation, supervision, project administration, and funding acquisition. All authors have agreed on their respective contributions and approved the final version of the manuscript. The corresponding author is responsible for handling all communications related to submission, revision, and publication processes, and ensures that all listed authors are kept informed throughout these stages.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Fatayat	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Wahyu Lestari		✓				✓		✓	✓	✓	✓	✓		
Alfirman	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

In this study, the authors declare that they have no known financial conflicts of interest or personal relationships that could have influenced the work reported in this paper. The authors declare no conflicts of interest.

INFORMED CONSENT

In situations where disclosure or identification of personal information is necessary for scientific purposes, the authors have obtained complete documentation of written informed consent, including written assent from individuals prior to their participation in the study. We hereby declare that we have obtained written informed consent from all individuals included in this study.

DATA AVAILABILITY

- In this study, the data are openly available in the Roboflow Universe repository at the following URL: https://universe.roboflow.com/ml-vpzcc/chili-plant-disease. Readers may access, view, and reuse this dataset according to the terms and conditions stated in the repository.

REFERENCES

[1] G. Wilkie, Object-oriented software engineering. *Addison Wesley Publishing Company*, 1993.

[2] D. H. Levine, "Introduction to Neural and Cognitive Modeling," *Psychology Press*, 2000. doi: 10.4324/9781410605504.

[3] R. S. Ledley, "Building Microsoft SQL Server 7 Applications with COM; By Sanjiv Purba. John Wiley & Sons, Inc. ISBN: 0-471-19233-3.," *Computer Languages*, vol. 25, no. 2, pp. 114–115, Jul. 1999, doi: 10.1016/s0096-0551(99)00010-7.

[4] M. I. Kamil and A. P. Wibowo, "Implementation of the Naive Bayes classifier algorithm for classifying toddler nutritional status," *Journal of Applied Informatics and Computing*, vol. 8, no. 2, pp. 421–427, Nov. 2024, doi: 10.30871/jaic.v8i2.8669.

[5] K. D. Sari and A. Eviyanti, "Expert system for diagnosing human skin diseases using web-based naïve bayes method," Procedia of Engineering and Life Science, vol. 1, no. 2, Jul. 2021, doi: 10.21070/pels.v1i2.1021.

[6] R. Irsyada, N. Cahyani, and L. Badriyah, "Innovation of an expert system for diagnosing allergic diseases in children using the web-based certainty factor method," *Brilliance: Research of Artificial Intelligence*, vol. 4, no. 2, pp. 909–918, Jan. 2025, doi: 10.47709/brilliance.v4i2.5204.

[7] W. Setiawan and S. Ratnasari, "Expert system for diagnosis of eye diseases using naive bayes classifier (In Indonesia)," in *Proceedings of Semnastek 2014*, 2014, pp. 1–6. [Online]. Available: https://jurnal.umj.ac.id/index.php/semnastek/article/view/345

[8] H. T. Sihotang, "Expert system for diagnosing diseases in corn plants using the Bayesian method (In Indonesia)." Aug. 25, 2019. doi: 10.31227/osf.io/dguhb.

[9] A. S. R. Sinaga and D. Simanjuntak, "Expert system for detecting malnutrition in toddlers using the Naïve Bayes classifier method (In Indonesia)," *Jurnal Inkofar*, vol. 1, no. 2, Jan. 2020, doi: 10.46846/jurnalinkofar.v1i2.110.

[10] S. Law, A.W Leung, and C. Xu, "Severe acute respiratory syndrome (SARS) and coronavirus disease-2019 (COVID-19): From causes to preventions in Hong Kong," *International Journal of Infectious Diseases*, vol. 94, pp. 156–163, 2020, doi: 10.1016/j.ijid.2020.03.059.

[11] A. Rahman and D. Wulandari, "Implementation of the Naïve Bayes algorithm in assigning academic advisors for university students," *Journal of Information Systems and Technology*, vol. 4, no. 3, 189–197, 2002.

[12] H. Zhang, "The optimality of Naive Bayes," In *Proceedings of the 17th International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, vol. 1, no. 2, 2004.

[13] M. Fadlillah, E. Mulyana, and D. Nugraha, "Application of the Naïve Bayes algorithm for determining thesis supervisors in an academic information system." *Journal of Technology and Information Systems*, vol. 6, no. 2, pp. 105–112, 2020.

[14] Y. Yanti and S. Sulindawaty, "Expert system diagnose disease persian cat with Bayes theorem method," *Journal Of Computer Networks, Architecture and High Performance Computing*, vol. 2, no. 1, pp. 149–157, Jan. 2020, doi: 10.47709/cnapc.v2i1.380.

[15] Y. Oktavia, I. Kanedi, and S. Sapri, "Application of the Naive Bayes method in diagnosing diseases of garlic plants," *Jurnal Komputer Indonesia*, vol. 1, no. 2, Dec. 2022, doi: 10.37676/jki.v1i2.22.

[16] S. Susanti and O. Manahan, "Disease diagnosis expert system at chili plants using bayes method," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 2, no. 2, pp. 292–296, 2020, doi: 10.47709/cnapc.v2i2.432.

[17] D. Syahfitri, B. S. Ginting, and A. Sihombing, "Expert system to diagnose pregnancy diseases in women using naive bayes method," *Journal of Information Systems and Technology Research*, vol. 1, no. 2, pp. 98–102, May 2022, doi: 10.55537/jistr.v1i2.148.

[18] M. Karmila, E. Maria, and A. Franz, "Expert system for diagnosis of pepper plant diseases using certainty factor and Naïve Bayes Methods," *Tepian*, vol. 2, no. 4, pp. 174-181, 2021, doi: 10.51967/tepian.v2i4.744.

[19] L. P. Wanti, , N. O. Fiarni, , and M. Handayani, "Optimation certainty factor method for an expert system to determinination large red Chili diseases," in *Proceedings of the 5th International Conference on Applied Science and Technology on Engineering Science (iCAST-ES 2022)*, pp. 11-19, doi: 10.5220/0011710500003575.

[20] H. N. Ngugi, A. A. Akinyelu, and A. E. Ezugwu, "Machine Learning and Deep Learning for Crop Disease Diagnosis: Performance Analysis and Review," *Agronomy*, vol. 14, no. 12, p. 3001, 2024, doi: 10.3390/agronomy14123001.

[21] L. Wan *et al*., "Identification of pepper leaf diseases based on TPSAO-AMWNet," *Plants*, vol. 13, no. 11, p. 1581, 2024, doi: 10.3390/plants13111581.

[22] I. Rish, "An empirical study of the Naive Bayes classifier," in *Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, 2001.

[23] A. Chen and M. Wasif, "Naïve Bayes classifier for text classification," *International Journal of Computer Applications*, vol. 182, no. 48, pp. 10–14, 2019.

[24] Y. Resti, C. Irsan, A. Neardiaty, C. Annabila, and I. Yani, "Fuzzy Discretization on the multinomial Naïve Bayes method for modeling multiclass classification of corn plant diseases and pests," *Mathematics*, vol. 11, no. 8, p. 1761, 2023.

BIOGRAPHIES OF AUTHORS

Fatayat, S.Kom, M.Kom. is a lecturer at Riau University, Indonesia. She obtained her Master's degree in Computer and Information Science from University Putra Indonesia (YPTK, Indonesia). He is currently pursuing her Ph.D. in computer science. His research interests include data mining and machine learning. She can be contacted at email: fatayat@lecturer.unri.ac.id.

Dra. Wahyu Lestari, M.Si., is a permanent lecturer at S1 Biology Study Programme, Faculty of Mathematics and Natural Sciences, Riau University. His last education was in Postgraduate (S2) Biology (FMIPA) at. the Bandung Institute of Technology. Research studies related to plant physiology and plant ecophysiology. She can be contacted at email: wahyu.lestari@lecturer.unri.ac.id.

Alfirman, S.Kom, M.Kom. is a lecturer at Riau University, Indonesia. He obtained his Master's degree in Computer and Information Science from Universitas Putra Indonesia (YPTK, Indonesia). He is currently pursuing his Ph.D. in computer science. His research interests include data mining and machine learning. He can be contacted at email: alfirman@lecturer.unri.ac.id.