RESEARCH ARTICLE OPEN ACCESS

Comparison of Naïve Bayes and Certainty Factor Method for Corn Disease Expert System: Case in Bangkalan, Indonesia

Mohammad Syarief, Novi Prastiti, Wahyudi Setiawan
(Department of Informatics Management, Faculty of Engineering, Trunojoyo University, Bangkalan, Indonesia
Coresponding author: wsetiawan@trunojoyo.ac.id

ABSTRACT
One of the application of expert system is in agriculture. In this study the expert system is applied to detect the disease in corn plants. Symptoms that indicate a particular disease, serve as a knowledge base on the system. The symptoms used are 46 which can indicate one of 15 types of diseases in corn plants. There are two methods used to classify the diseases of Naïve Bayes and Certainty Factor. The Naïve Bayes method is a simple classfier type that applies Bayes theorems with strong (naïve) independent assumptions. While Certainty Factor is a Classifier method that has a variation in the level of confidence or weight on each of the symptoms entered. Based on the test results, the system using the Certainty Factor method has a better percentage of truth than using the Naïve Bayes method. Experiment result shows the accuracy of up to 80%.

Keywords: Naïve Bayes, Certainty Factor, Corn Pest, Corn Disease, Expert System.

Date of Submission: 31-10-2017 Date of acceptance: 09-11-2017

I. INTRODUCTION

The corn is a staple food for residents of the island of Madura. The condition of infertile soil caused by lack of water source makes the farmers look for alternative staple crops other than rice. Corn does not need as much water as rice. So it's easier to cultivate. But the cultivation of corn crops also need to pay attention to the factors that influence it. Factors that affect the growth of corn plants such as seed selection, planting methods, how to care and the prevention of pests and diseases.

Agricultural production obtained through learning of agricultural cultivation knowledge and knowledge about corn disease. This study provides an alternative solution to know the disease in corn plants, it is expected emerge various growth disorders such as occur crop failure can be avoided more leverage. There are several causes of diseases that arise in corn plants due to be fungal and viral bacteria. The growth of plants affected by the disease can be disrupted, for example, the leaves look pointy and small, the cob becomes deformed/dwarf and the leaves look to dry up. Factors that cause crop failure can be caused by the understanding that farmers do not know what types of diseases are attacking and do not know how to control when their plants are attacked by diseases.

.Generally disease detection is done by agricultural or expert personnel. Limitations on the number of experts in the field of agriculture can be given an alternative solution that is using expert systems. Expert system is a system that transfer knowledge from the real experts into a computerized system. One of the application of expert system is in agriculture. In this study expert systems are used to detect diseases or pests in corn plants.

II. LITERATURE REVIEW

2.1 Previous Research

There are several research on expert system of corn disease such as Research which use Ada-Boost algorithm to get high level of accuracy of the expert system created. Expert system created web-based with three types of users are End user / farmer, experts and admin. Knowledge based used in the form of disease or pest is detected as many as ten along with the symptoms [1].

Further research is about fuzzy expert system for diagnosis and treatment of corn disease. Expert systems have five symptoms with five linguistic variables: Very Low, Low, Medium, High and Very High [2].

Subsequent research is about expert systems to detect and diagnose leaf diseases in three types of plants namely rice, corn and wheat. The system is built using image processing and machine vision. The symptoms and diseases as part of the knowledge base are stored in a database. Web-based expert system applications [3].

Further research is about Expert System of disease diagnosis using image of corn plant. There are six classes of detected diseases: ear, stem, leaf, root, seed and whole plant. Knowledge base consists of 23 images of the type of disease. Expert system created web-based and named Corn Diseases Diagnosis (CDDES) [4].

These studies do not provide reports on the percentage of classification accuracy results. So in this study will be calculated the percentage of classification based on the experiments performed.

2.2 Naïve Bayes Classifier

Naïve Bayes Classifier is a simple classification method that uses Bayes's basic theorem with strong (Naïve) independent assumptions. Naïve Bayes calculations use equation (1) [5].

$$P\left(a_i \mid v_j\right) = \left(n_c + (m \times p)\right) / (n + p) \qquad (1)$$

$P\left(a_i \mid v_j\right)$ = chance of attributes (attributes) if known circumstances v_j

n_c = the same symptom value between learning and testing data (set 0 if none or set 1 if any)

m = number of symptoms

p = 1/ number of classes or diseases

n = 1

To calculate the highest probability of a case the following algorithm is used [6]:

1. Calculate value n_c . Value n_c is the same symptom data in data learning. Value n_c is a value of 0 and 1. The value is set 0 if no symptoms are found, the value is set 1 if symptoms are found in a case to be searched for the type of illness.
2. Determine the value of each $P\left(a_i \mid V_j\right)$
3. Calculate value $P\left(V_j\right)$
4. Multiply results $P\left(a_i \mid V_j\right)$ with $P\left(V_j\right)$ for each v
5. Determining the winner with the largest multiply result from step 4

2.3 Certainty Factor

Certainty Factor (CF) is a classification method that uses the value of confidence or weight for the symptoms experienced by an object. This is to overcome the uncertainty in decision making. There are two types of values in CF, namely [7]:

1. Assessment of an expert who sets the standard for the assessment of the symptoms of the disease
2. The value of the user to show the degree of confidence of the premise of symptoms of the disease. Usually done by users who have had previous experience.

The equation of Certainty Factor is as follows:

$$CF(H,E) = MB(H,E) - MD(H,E) \qquad (2)$$

CF (H, E) = Certainty Factor, with H and Evidence Hypothesis

MB (H, E) = Measure Believe, Confidence Level for H Hypothesis, if given Evidence E (between 0 and 1)

MD (H, E) = Measure Disbelief, Level of Uncertainty for H Hypothesis, if given Evidence E (between 0 and 1)

Certainty Factor to other conditions, can be shown in the following equation [8]:

Certainty Factor for a single premise:

$$CF(H,E) = CF(E) \times CF(rule)$$

$$CF(H,E) = CF(User) \times CF(Expert) \qquad (3)$$

Certainty Factor for multiple premises:

$$CF\ (\text{A or B}) = Min\left(CF(A), CF(B)\right) \times CF(rule)$$

$$CF\ (\text{A and B}) = Max\left(CF(A), CF(B)\right) \times CF(rule) \qquad (4)$$

Certainty Factor for similarity rules:

CFcom (CF1,CF2)=

$$CF1 + CF2 - (CF1 \times CF2), \quad CF1 \text{ and } CF2 > 0$$

$$CF1 + CF2 \times (1 + CF1), \quad CF1 \text{ and } CF2 < 0$$

(CF1+CF2)/ 1- min(|CF1|,|CF2|) ,CF1 or CF2> 0

(5)

Cfcom = CF combinations,

CF1 = CF rule 1

CF2 = CF rule 2

III. EXPERIMENTAL RESEARCH

The data used as a test consisting of 46 symptoms and 15 kinds of diseases [9]. Tests were performed by calculating the probability of the type of illness from known symptoms. The following will be presented the first case example. Known symptoms in corn plants are:
g1 The presence of bite marks on leaves
g2 Withered leaves
g4 The existence of bite marks on the trunk
g20 Leaf bone is damaged
g21 Stem rot

By using Naïve Bayes algorithm obtained calculation in Table 1

Table 1. Calculation results of Naïve Bayes

No	Pest/Disease	p					Value of v	Grade
		1	2	4	20	21		
1	Seeds Flies	0.0865248	0.0865248	0.0652482	0.0652482	0.0865248	1.8385E-07	1
2	Soil Caterpillar	0.0652482	0.0652482	0.0865248	0.0652482	0.0652482	1.0455E-07	2
3	Lundi (uret)	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
4	Bulai	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
5	Dwarf Mozaic Virus	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
6	Stem Borer	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
7	Spot Leaf	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
8	Hawar/Upih Leaf	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
9	Grayak Caterpillar	0.0652482	0.0652482	0.0652482	0.0865248	0.0652482	1.0455E-07	2
10	Rotten Stem and Cobs	0.0652482	0.0652482	0.0652482	0.0652482	0.0865248	1.0455E-07	2
11	Rust	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
12	Burned and swollen	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
13	Grasshoper	0.0865248	0.0652482	0.0865248	0.0865248	0.0652482	1.8385E-07	1
14	Corncob borer	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3
15	Grasshopper Corn	0.0652482	0.0652482	0.0652482	0.0652482	0.0652482	7.8841E-08	3

In the example of this classification with unknown symptoms 1,2,4,20,21 produce the highest probability value Flies seeds and grasshoppers. There are two possible types of diseases with the same probability values. Using the same case example. Disease detection problems can be solved with certainty factor method, can be completed using the following steps:

1. Determine the weight or level of confidence of any symptoms that cause a particular disease. The assessment range given by experts is shown in table 2.

Table 2. Confidence level of the symptoms thacause the disease

No	Value	Meaning of CF
1	0.2	Not Sure
2	0.4	Little Sure
3	0.6	Sure
4	0.8	Very Sure

2. Determine the highest certainty factor of each type of pests and diseases. Naïve Bayes method of calculation obtained the highest probability value contained in seed flies and grasshoppers. Next is calculated using CF. Given the symptoms found in corn plants is a symptom no 1,2,4,20,21. Calculation CF using equation (5)

Table 3. Certainty Factor of pest/Diasease

Pest/Disease	No. of symptoms	Certainty Factor of the symptoms				
		1	2	4	20	21
Seed flies	1,2,21	0.6	0.8	-	-	0.2
grasshoppers	1,4,20	0.8	-	0.4	0.6	-

Table 4. Certainty Factor of All Pest/Disease

No	Pest/Disease	CF1	CF2	CF3	CF
1	Seeds Flies	0.6	0.8	0.2	0.936
2	Soil Caterpillar	0.4	0	0	0.4
3	Lundi (uret)	0	0	0	0
4	Bulai	0	0	0	0
5	Dwarf Mozaic Virus	0	0	0	0
6	Stem Borer	0	0	0	0
7	Spot Leaf	0	0	0	0
8	Hawar/Upih Leaf	0	0	0	0
9	Grayak Caterpillar	0.6	0	0	0.6
10	Rotten Stem and Cobs	0.8	0	0	0.8
11	Rust	0	0	0	0
12	Burned and swollen	0	0	0	0
13	Grasshoper	0.8	0.4	0.6	0.952
14	Corncob borer	0	0	0	0
15	Grasshopper Corn	0.4	0	0	0.4

Based on Table 4. Symptoms 1,2,4,20 and 21 pest / disease Certainty Factor Grasshopper with the greatest value is 0,952

IV. RESULT

Tests based on 15 maize plants using methods Certainty Factor and validation of experts obtained the results shown in Table 5.

Table 5.Experiment Result

Corn	Expert System Detection	Expert opinion
1	Grayak	Grayak
2	Grasshopper	Grasshopper
3	Grayak	Grayak
4	Seeds Flies	Seeds Flies
5	Bulai	Bulai
6	Grayak	Grayak
7	Grayak	Grayak
8	Grasshopper	Grasshopper
9	Stem Borer	Grayak
10	Grayak	Grayak
11	Bulai	Stem Borer
12	Soil Caterpillar	Mozaik Virus
13	Bulai	Bulai
14	Grayak	Grayak
15	Seeds Flies	Seeds Flies

From Table 5 , The Accuracy = correct prediction / overall observation x 100% = 12/15 x 100% = 80% The Interface of the expert system is shown in Figure 1 and 2.

Fig 1. The first Interface of Expert System

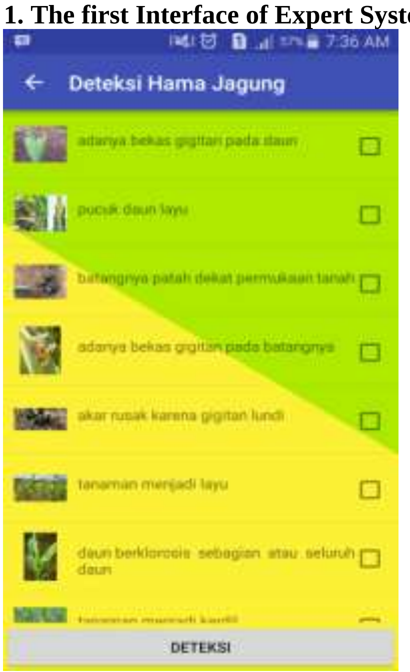

Fig 2. Pest-Disease of Corn Detection

V. CONCLUSIONS

1. Method of Certainty Factor to classify the type of pest / disease better than the Naïve Bayes method . For the case study in the study using 15 cases, the percentage of classification up to 80%
2. Research can be developed to develop a type of disease during storage of corn in the barn. The views of other experts who can be used to perform a variation on pest and disease detection

ACKNOWLEDGMENT

We would like to thank the research team to the Director General of Research and Community Service DRPM DIKTI who has funded this Research on Research Scheme Penelitian Dosen Pemula. We would also like to thank to Dr. Achmad Amzeri who has become an expert in this research.

REFERENCES

[1]. P. Kumar, R. Babu, and G. Kumar, "Maize Expert System", Computer Science & Engineering : An International Journal (CSEIJ), vol 2 no 3, June 2012, 43-55.

[2]. O.C. Agbonifo, and D.B. Olufolaji, "A Fuzzy Expert System for Diagnosis and Treatment of Maize Plant Diseases", International Journal of Advanced Research in Computer Science and Software Engineering, vol 2 no 3, December 2012, 83-89.

[3]. R. Kaur, S. Din, and P. Pannu, "Expert System to Detect and Diagnose the Leaf Diseases of Cereals", International Journal of Current Engineering and Technology, vol 3 no 4, October 2013, 1480 – 1483.

[4]. L. Jun-Chen, M. Bo, L. Shao-kun, W. Ke-ru , and G. Shi-ju, " AN Image Based Diagnostic Expert System for Corn Diseases", Agricultural Science in China, vol 9 no 8, August 2010, 1221-1229

[5]. D. C. Corrales, J. C. Corrales and A. Figueroa-Casas, "Towards detecting crop diseases and pest by supervised learning", Ing. Univ., vol. 19, no. 1, pp. 207-228, Ene.-Jun., 2015.

[6]. W. Setiawan, & S. Ratnasari, "Sistem Pakar Diagnosis Penyakit Mata Menggunakan Naïve Bayes Classifier". *Prosiding Seminar Nasional Sains Dan Teknologi*, Universitas Muhammadiyah Jakarta, 2014

[7]. I. Hatzilygeroudis, A. Karatrantou, and C. Pierrakeas, "PASS: An Expert System with Certainty Factors for Predicting Student Success", 2004, LNAI 3213, pp. 292–298.

[8]. T. Mellouli, Complex Certainty Factors for Rule Based Systems – Detecting Inconsistent Argumentations, http:// ceur-ws.org/Vol-1335/wlp2014_paper6.pdf

[9]. M. Syarief, A. Mukminin, N. Prastiti, W. Setiawan, Penerapan Metode Naïve Bayes Classifier Untuk Deteksi Penyakit Pada Tanaman Jagung, NERO, Vol. 3 No. 1, 2017, 61- 68

Mohammad Syarief Comparison of Naïve Bayes and Certainty Factor Method for Corn Disease Expert System: Case in Bangkalan, Indonesia." International Journal of Engineering Research and Applications (IJERA) , vol. 7, no. 11, 2017, pp. 30-34.