Expert System of Diagnosing Disease in Banana Plants Using Dempster-Shafer and Certainty Factor

1st Freza Riana
Teknik Informatika
Ibn Khaldun University
Jl. K.H. Soleh Iskandar Km 2 Bogor
freza@ft.uika-bogor.ac.id

2nd Fitrah Satrya Fajar Kusumah
Teknik Informatika
Ibn Khaldun University
Jl. K.H. Soleh Iskandar Km 2 Bogor

Abstract—**The increasing demand for banana production for local and export needs is followed by the increasing need for quality banana plant seeds. However, the growth of banana plants is always disrupted by attacks by pests and diseases, especially *Fusarium Wilt* disease, so that this is one of the factors affecting the quality of banana plants and can be a threat to the decreasing quantity of banana production. The ignorance of banana planters and still at least a diagnosis of banana plant diseases is very important to make an inference engine in the form of a banana plant expert system. The Dempster-Shafer method and Certainty Factor are proposed to be applied to expert systems. The appropriateness of the Dempster-Shafer simulation simulation results with expert answers can be a clue that Dempster-Shafer and Certainty Factor can be used as an inference engine to diagnose banana plant diseases.**

Keywords: Dempster-Shafer, Certainty Factor, expert system, Fusarium Wilt, banana plant disease

I. INTRODUCTION

Banana plants are currently cultivated throughout the world, in the tropics and subtropics. The main production zones are located in Asia, representing 44% of world volume, followed by Africa (25%) and Central and South America (22%). Indonesia is ranked 8th as the country's center of world banana production [1]. Banana plants have the potential to be developed on an agribusiness scale, but their development is affected by pests and diseases [2]. The Food and Agriculture Organization of the United Nations (FAO) reported in 2015 that global banana exports excluding Raja Banana, decreased by 18.6 million tons in 2014. The decrease was due to the adverse effects of El Nino weather and the spread of Fusarium Wilt disease, resulting in production shortages in various countries banana production centers [3]. Indonesia also experienced the same thing, the development of exports over the past 13 years (1989-2013) has continued to decline due to the attacks of Fusarium Wilt and Bacterial Wilt which resulted in several Indonesian banana plantations stopping their production processes [4]. The United Nations Conference on Trade and Development reported that the most dangerous diseases in banana plantations caused by pathogenic fungi were the Sigatoka Spots and Fusrium Wilt [3]. Some other pathogenic fungi that cause disease in banana plants are Cordona Leaf Spots, Cross-Spot Spots [5].

Symptoms that cause diseases in banana plants are not yet familiar to banana farmers in Indonesia. Besides the diagnosis of diseases in banana plants has not been done much [6]. So, the ignorance of banana growers with this, can

threaten the production of bananas, both in terms of quality and quantity [7]. So that this does not happen, we need an inference machine to help diagnose banana plant diseases, namely an expert system. An expert system is a computer program that contains knowledge from one or more human experts about a specific field. Dempster Shafer (D-S) and Certainty Factor are the two methods proposed to be developed as assistant expert systems. D-S theory [8] is a mathematical theory, which allows a person to combine evidence from various sources and arrive at the trust function by taking into account all available evidence. While CF is a method to prove whether a fact is something that is certain or uncertain [9].

Studies that have implemented DS are detecting structural damage by combining results from data to reduce uncertainty caused by noise and model error measurements [10], diagnosis of Indihome service interruptions [11], The prediction for listed companies financial distress by using multiple prediction methods with rough set and D–S [12], conflicy-management in D-S using the degree of fasilty[13], exploring the combination of D-S theory and neural network for predicting trust and distrust [14] and diagnosis of acute respiratory infections, by combining the method of Integrating Variable-Centered Intelligent Rule System [15]. From some of these studies, the D-S Method can be proposed to be used in finding opportunities or possible trends. In addition to the D-S method, CF is also used in a number of studies that are referred to, expert system of Catfish Disease [16], detecting inconsistent argumentation [17], diagnosis red chili peppers diseases [18] and comparing D-S and CF for diagnosing stroke [19].

This research is a basic research to collect every symptom in every banana plant disease through a banana plant expert. The D-S and CF theories have been proposed to diagnose banana disease from the symptoms found. With this research, it is hoped that the inference engine of a banana disease expert system can be a quick solution to help diagnose diseases in banana plants.

II. METHOD

This research starts from a literture study of the symptoms and diseases of banana plants, then the acquisition of knowledge with a banana plant expert who works at the Lembang West Java Vegetable Research Center. Experts provide information in the form of data symptoms and names of diseases that attack banana plants. Furthermore, every

symptom in every disease is given a mass function by an expert. The data used in this study is to simulate several cases. The number of diseases as much as 4, namely: Sigatoga leaf spot, Cordona leaf spot, cross-leaf spot and fusarium wilt. The number of symptoms as many as 20 symptoms.

A. Dempster-Shafer

Every disease has a number of evidence that will be used on the uncertainty factor in making decisions to diagnose a disease. To overcome this number of evidence, Dempster-Shafer Theory uses rules that are better known as *Dempster's Rule of Combination* [20],[21]:

$$m_3(Z) = \sum_{X \cap Y = Z} m_1(X) m_2(Y) \qquad (1)$$

With m_3 *(Z)*: *mass function* dari *evidence* (Z), m_1 *(X)*: *mass function* dari *evidence* (X), m_2 *(Y)*: *mass function* dari *evidence* (Y). The mass function of the level of confidence of an evidence (symptom) is called an evidence measure. If the fact of the results obtained $\neq 1$ must be normalized from each element of the set where each element will be divided by:

$$1 - K \qquad (2)$$

$$K = \sum_{X \cap Y = \emptyset} m_1(X) m_2(Y) \qquad (3)$$

Where K is the variable that produces the zero set $\emptyset$. So that equations (1), (2) and (3) are generally substituted produce

$$m_3(Z) = \frac{\sum_{X \cap Y = Z} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y \neq \emptyset} m_1(X) m_2(Y)} \qquad (4)$$

B. Certainty Factor

The CF method shows a measure of certainty about a fact or rule. CF is the clinical parameter value given by MYCIN to show the amount of trust. The advantage of the CF method is that it can measure something that is certain or uncertain in decision making in an expert system of disease diagnosis. The basic formula CF [22]:

$$(h, e) = (h, e) - (h, e)$$
$$(4)$$

The next calculation is a combination of two or more rules with different evidence but in the same hypothesis:
$$\textbf{Rule 1} \quad CF(\textbf{h}, \textbf{e}_1) = CF_1 = C(\textbf{e}_1) \times CF \ (\textbf{Rule 1}) \quad (5)$$
$$Rule\ 2 \quad (h, e_2) = \textbf{CF}_2 = (\textbf{e}_2) \times (Rule\ 2) \quad (6)$$
$$CFkomb\ (\textbf{CF}_1, \textbf{CF}_2)) = \textbf{CF}_1 + \textbf{CF}_2(1 - \textbf{CF}_1) \qquad (7)$$

III. RESULT AND DISCUSSION

The results of testing the system with experts have been done and get the results that the system can diagnose fusarium wilt disease and other banana plant diseases with a good degree of accuracy. How the expert system works is to help laypeople be able to consult the system like consulting an expert. This expert system of diagnosing banana plant diseases will receive input on symptoms seen in banana plants by users. Testing is done by comparing the results of the diagnosis between the answers of the experts and the

expert system of diagnosing banana diseases. Testing of the system is done using 30 test data resulting in 100% accuracy.

Plant diseases detected by the inference machine using the D-S and CF methods produce the same disease, but the confidence value or percentage of confidence is different. This can be seen from one of the existing case simulations:

A. *Application of Dempster Shafer Calculation in the Determination of Diseases In Banana Plants Six Symptoms are Found in Banana Plants:*

- There is a change in the colour of the vessel tissue to brown on the pseudo stem; in the form of brown dots when pseudo stems are cut across or brown lines extending when pseudo stems are cut longitudinally
- There is necrosis in the tubers. When the weevil is cut across, there is brown to black necrosis around the tuber
- Parallel blackish brown lines on the leaf blade
- Elongated patches measuring 2 x 20 mm parallel to the leaf blade
- Pseudo stem split or broken
- At the beginning of development, the spots are arranged in line with the mother's leaf bone (midrib)

1) There is a change in the color of the vessel tissue to brown on the pseudo stem. In the form of brown dots when pseudo stems are cut across or brown lines extending when pseudo stems are cut longitudinally, $m_1(P_1) = 0.9, m_1(\Theta) = 1 - m_1(P_1) = 1 - 0.9 = 0.1$.

2) There is necrosis in the tubers. When the weevil is cut across, there is brown to black necrosis around the tuber, $m_2(P_1) = 0.9, \ m_2(\Theta) = 0.1$

TABLE 1. COMBINATION OF MASS FUNCTION m_1 AND m_2

	$m_2(P_1) = 0.9$	$m_2(\Theta) = 0.1$
$m_1(P_1) = 0.9$	$(P_1) = 0.81$	$(P_1) = 0.9$
$m_1(\Theta) = 0.1$	$(P_1) = 0.9$	$\Theta = 0.001$

$$m_3(P_1) = \frac{0.81 + 0.9 + 0.9}{1 - 0} = 0.99$$
$$m_3(\Theta) = \frac{0.001}{1 - 0} = 0.001$$

3) Parallel blackish brown lines on the leaf blade

TABLE 2. COMBINATION OF MASS FUNCTION m_3 AND m_4

	$m_4(P_2) = 0.9$	$m_4(\Theta) = 0.1$
$m_3(P_1) = 0.99$	$\Theta = 0.891$	$(P_1) = 0.099$
$m_3(\Theta) = .01$	$(P_2) = 0.009$	$\Theta = 0.001$

$$m_5(P_1) = \frac{0.099}{1 - 0.891} = 0.908527$$
$$m_5(P_2) = \frac{0.009}{1 - 0.891} = 0.082569$$
$$m_5(\Theta) = \frac{0.001}{1 - 0.891} = 0.009174$$

4) Elongated patches measuring 2 x 20 mm parallel to the leaf blade

TABLE 3. COMBINATION OF MASS FUNCTION m_5 AND m_6

	$m_6(P_2) = 0.85$	$m_6(\Theta) = 0.15$
$m_5(P_1) = 0.908527$	$\emptyset = 0.772018$	$(P_1) = 0.136239$
$m_5(P_2) = 0.082569$	$(P_2) = 0.070183$	$(P_2) = 0.012385$
$m_5(\Theta) = 0.009174$	$(P_2) = 0.007798$	$\Theta = 0.001376$

$$m_7(P_1) = \frac{0.136239}{1 - 0.772018} = 0.597586$$

$$m_7(P_2) = \frac{0.070183 + 0.007798 + 0.012385}{1 - 0.772018}$$
$$= 0.396378$$

$$m_7(\Theta) = \frac{0.001376}{1 - 0.772018} = 0.006036$$

5) Pseudo stem split or broken

TABLE 4. COMBINATION OF MASS FUNCTION m_7 AND m_8

	$m_8(P_1) = 0.8$	$\Theta = 0.2$
$m_7(P_1) = 0.597586$	$(P_1) = 0.478068$	$(P_1) = 0.07926$
$m_7(P_2) = 0.396378$	$\emptyset = 0.317103$	$(P_2) = 0.079276$
$\Theta = 0.006036$	$(P_1) = 0.004829$	$\Theta = 0.001207$

$$m_9(P_1) = \frac{0.478068 + 0.07926 + 0.004829}{1 - 0.317103}$$
$$= \mathbf{0.882145}$$

$$m_9(P_2) = \frac{0.079276}{1 - 0.317103} = \mathbf{0.116087}$$

$$m_9(\Theta) = \frac{0.001207}{1 - 0.317103} = 0.001768$$

B. Application of CF Calculation in the determination of diseases in banana plants

1) The possibility of banana plants affected by fusarium wilt

- There is a change in the color of the vessel tissue to brown on the pseudo stem; in the form of brown dots when pseudo stems are cut across or brown lines extending when pseudo stems are cut longitudinally
- There is necrosis in the tubers. When the weevil is cut across, there is brown to black necrosis around the tuber
- Pseudo stem split or broken

TABLE 5. COMBINATION OF RULE RESULTS FOR EACH SYMPTOM IN FUSARIUM WILT

Trust of Values from Expert	Trust of Values from User	TE*TU	$CF_1 + CF_2 * (1 - CF_1)$
0.9	0.8	0.72	0.72
0.9	0.8	0.72	0.9216
0.8	0.8	0.64	**0.971776**

2) 3.2.2 The possibility of banana plants affected Sigatoga Spaces

- Parallel blackish brown lines on the leaf blade
- Elongated patches measuring 2 x 20 mm parallel to the leaf blade
- At the beginning of development, the spots are arranged in line with the mother's leaf bone (midrib)

TABLE 6. COMBINATION OF RULE RESULTS FOR EACH SYMPTOM IN SIGATOGA LEAF SPOT

Trust of Values from Expert	Trust of Values from User	TE*TU	$CF_1 + CF_2 * (1 - CF_1)$
0.9	0.8	0.72	0.72
0.85	0.8	0.68	0.9104

The results of the D-S and CF calculations show that of the six symptoms that have been found, namely fusarium wilt disease. Suitability results are the same as those obtained with expert answers.

IV. CONCLUSION

Both the D-S and CF methods can work well in the inference system, this is indicated by the results of simulation answers from experts and the results of both the D-S and CF methods are the same. Plant diseases detected by the inference machine using the D-S and CF methods produce the same disease, but the confidence value or percentage of confidence is different. System Calculation and Manual Calculation use the Dempster Shafer Method, if there is a difference in the results of the trust value, it is caused because the system calculates the high trust value first.

System calculations and manual calculations use the Certainty Factor, producing the same trust value. But the Certainty Factor method is very influential on input values from users

REFERENCES

[1] United Nation Conference On Trade And Development. *Banana*. Newyork and Geeneva. 2016

[2] C. Hermanto, J. Jumjunidang, R. P. Yanda, and N. Nasir, "Uji Virulensi Isolat Fusarium oxysporum f.sp. cubense Dalam Vegetative Compatibility Group Complex 0124 Pada Tanaman Pisang," *J.Hort*, vol. 23, no. 4, p. 372, May 2016.

[3] CIRAD Banana Diseases and Pests. Fruitrop.2015 (231):85-91.

[4] Kukuh Dwi Saputro, Made Dwi Setyadhi Mustika," vol. 4, no. 8, p. 28, 2015

[5] L.Soesanto, E. Mugiastuti, F. Achamd, Witcaksono. Diagnosis Lima Penyakit Karena Jamur Pada 100 Kultivar Bibit Tanaman Pisang, "*J.Hort*, vol.12, n0.1, p 36-45, Maret 2012.

[6] D. Emilda and M. Istianto, "Pengaruh Minyak Cengkeh terhadap Pertumbuhan Koloni dan Sifat Antagonis Cendawan Gliocladium sp. terhadap Fusarium oxysporum f. sp. cubense," *J.Hort*, vol. 21, no. 1, p. 33, Oct. 2016.

[7] J. Hort, "Penyakit Layu Fusarium Pada Tanaman Pisang di Provinsi," vol. 22, no. 2, p. 8, 2012.

[8] Shafer G. A mathematical theory of evidence. Princenton NJ: Princenton University Press. 1976

[9] Huadong Wu, M. Siegel, R. Stiefelhagen, and Jie Yang, "Sensor fusion using Dempster-Shafer theory [for context-aware HCI]," in *IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No.00CH37276)*,

Anchorage, AK, USA, 2002, vol. 1, pp. 7–12.

[10] Y. Bao, H. Li, Y. An, and J. Ou, "Dempster–Shafer evidence theory approach to structural damage detection," *Structural Health Monitoring*, vol. 11, no. 1, pp. 13–26, Jan. 2012.

[11] E. Lestari and E. U. Artha, "Sistem Pakar dengan Metode Dempster Shafer untuk Diagnosis Gangguan Layanan INDIHOME di PT TELKOM Magelang," *KHIF*, vol. 3, no. 1, p. 16, Jul. 2017.

[12] Zhi Xiao, Xianglei Yang, Ying Pang, and Xin Dang, "The prediction for listed companies' financial distress by using multiple prediction methods with rough set and Dempster–Shafer evidence theory," *Knowledge-Based Systems*, vol. 26, pp. 196–206, 2012.

[13] J. Schubert, "Conflict management in Dempster–Shafer theory using the degree of falsity," *International Journal of Approximate Reasoning*, vol. 52, no. 3, pp. 449–460, Mar. 2011..

[14] X. Wang, Y. Wang, and H. Sun, "Exploring the Combination of Dempster-Shafer Theory and Neural Network for Predicting Trust and Distrust," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–12, 2016.

[15] S. A. S. Mola, N. D. Rumlaklak, and N. P. D. Prityaningsih, "Integrasi Variable-Centered Intelligent Rule System dengan Teori Dempster-Shafer pada Sistem Pakar Infeksi Saluran Pernafasan Akut," *J. Sistem Info. Bisnis*, vol. 9, no. 1, p. 71, May 2019.

[16] I. Sumarono, D. Arisandy, A.Siahaan, and Mesran. "Expert System of Catfish Disease Determinants Using Certainty Factor Method," *IJRTER*, vol. 3, no. 8, pp. 202–209, Aug. 2017.

[17] T. Mellouli, "Complex Certainty Factors for Rule Based Systems – Detecting Inconsistent Argumentations," p. 22.

[18] Br Sembiring NS, Sinaga MD 2018 Penerapan Metode Dempster Shafer Untuk Mendiagnosa Penyakit dari Akibat Bakteri Treponema Pallidum. CSRID. Mar 7;9(3):134.

[19] E. K. Panggabean, "Comparative Analysis Of Dempster Shafer Method With Certainty Factor Method For Diagnose Stroke Diseases," *Int. J. Art. Intell. Research*, vol. 2, no. 1, p. 32, Mar. 2018.

[20] R. R. Yager and L. Liu, Eds., *Classic Works of the Dempster-Shafer Theory of Belief Functions*, vol. 219. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[21] Q. Chen, A. Whitbrook, U. Aickelin, and C. Roadknight, "Data classification using the Dempster–Shafer method," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 26, no. 4, pp. 493–517, Oct. 2014.

[22] Jacson, and Petter. Introduction to expert systems. Addison Wesley Longman Limited. 1999.